

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 22, 2003, 17:13:34 ; Search time 7709 Seconds
(without alignments)
11414.789 Million cell updates/sec

Title: US-09-601-168b-1

Perfect score: 2151

Sequence: 1 tgcgttgctgcgcctggc.....gtttgccagaaaaaaaaa 2151

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 288711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sta.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sta.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_hgt_hum.*
- 31: em_hgt_inv.*
- 32: em_hgt_other.*
- 33: em_hgt_mus.*
- 34: em_hgt_pln.*
- 35: em_hgt_rod.*
- 36: em_hgt_mam.*
- 37: em_hgt_vrt.*
- 38: em_sy.*
- 39: em_hgt_hum.*
- 40: em_hgt_mus.*
- 41: em_hgt_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2151	100.0	2151	6	AX019507	AX019507 Sequence
2	2151	100.0	2151	6	AX057166	AX057166 Sequence
3	2151	100.0	2151	6	BD131778	BD131778 Human bet
4	2151	100.0	2151	9	AF129530	AF129530 Homo sapi
5	2151	100.0	2151	9	HSBTRCP	Y14153 Homo sapien
6	2030.4	94.4	2308	9	BC027994	BC027994 Homo sapi
7	1682.2	78.2	2175	10	AF081887	AF081887 Mus muscu
8	1663.2	77.3	1818	9	AF101784	AF101784 Homo sapi
9	1630.4	75.8	2081	10	BC003989	BC003989 Mus muscu
10	1567.2	72.9	1979	10	AF110396	AF110396 Mus muscu
11	1451.2	67.5	1712	10	AF099932	AF099932 Mus muscu
12	1442.8	67.1	1710	10	AF112979	AF112979 Mus muscu
13	1017.8	47.3	1671	5	XELSBTRCP	M98468 African cla
14	883.4	41.1	2134	9	AB033279	AB033279 Homo sapi
15	883.4	41.1	2252	9	AB033280	AB033280 Homo sapi
16	883.4	41.1	2274	9	AB033281	AB033281 Homo sapi
17	883.4	41.1	4230	9	AB014596	AB014596 Homo sapi
18	883.4	41.1	4395	9	BC026213	BC026213 Homo sapi
19	882.4	41.0	1653	9	AF176022	AF176022 Homo sapi
20	877.8	40.8	3874	10	AB093260	AB093260 Mus muscu
21	877.8	40.8	4030	10	BC034261	BC034261 Mus muscu
22	853.2	39.7	1970	10	AY038079	AY038079 Mus muscu
23	810.2	37.7	2530	5	BC045356	BC045356 Danilo rer
24	757.8	35.2	67008	9	AP000252	AP000252 Homo sapi
25	757.8	35.2	100000	9	AP000212	AP000212 Homo sapi
26	757.8	35.2	149298	9	AP000031	AP000031 Homo sapi
27	757.8	35.2	340000	9	AP001711	AP001711 Homo sapi
28	742.6	34.5	2445	3	AB076893	AB076893 Clona int
29	742.6	34.5	2514	3	AK114399	AK114399 Clona int
30	676.4	31.4	2534	3	AV118898	AV118898 Drosophil
31	673.6	31.3	292639	2	AC112602	AC112602 Rattus no
32	673.2	31.3	2154	3	AF032878	AF032878 Drosophil
33	673.2	31.3	2367	3	AF222924	AF222924 Drosophil
34	673.2	31.3	3532	3	AF222923	AF222923 Drosophil
35	657	30.5	657	6	AX050273	AX050273 Sequence
36	518.8	24.1	1561	6	AX057148	AX057148 Sequence
37	512	23.8	934	5	XL063921	U63921 Xenopus lae
38	502.6	23.4	1827	5	XL063922	U63922 Xenopus lae
39	463.6	21.6	2983	10	BC008552	BC008552 Mus muscu
40	454	21.1	1443	3	AF339101	AF339101 Heteroder
41	453.2	21.1	2700	3	AF275253	AF275253 Caenorhab
42	386.2	18.0	1217	3	AK115763	AK115763 Clona int
43	333.2	15.5	108503	9	AL627424	AL627424 Human DNA
44	288	13.4	173634	3	AC009344	AC009344 Drosophil
45	288	13.4	235707	3	AE003733	AE003733 Drosophil

ALIGNMENTS

RESULT 1
AX019507
LOCUS AX019507 Sequence 1 from Patent WO9938969. 2151 bp DNA linear PAT 07-SEP-2000
DEFINITION AX019507
ACCESSION AX019507
VERSION AX019507.1 GI:110043427
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
Arenzana,S.F., Concordet,J.P., Kroll,M., Durand,H., Benarous,R. and Margottin,F.
AUTHORS Protein humaine beta -trcp
TITLE Patent: WO 9938969-A 1 05-AUG-1999;
JOURNAL

ARENZANA SEISDEDOS FERNANDO (FR); CONCORDET JEAN PAUL (FR); INST
NAT SANTE RECH MED (FR); KROLL MATHIAS (FR); DURAND HERVE (FR);
BENAROUS RICHARD (FR); MARGOTTIN FLORENCE (FR); PASTEUR INSTITUT
(FR)

FEATURES

source

Location/Qualifiers

1..2151
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="ADNC codant pour la proteine humaine BtrCp"
70..1779
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC07740.1"
/db_xref="GI:10043428"
/translation="MDPAEVLQEKALKFNWSSPERDQNGEPKKIIEKNSLRQTY
NSCARLCINQTVCLASTAMKTEVCYAKTKLANGTSSMIVPKORKLSAYEKEKLCV
KIFQWSSDOVEFVEHLISQMHYQHGINSYLKPLQDFITALPARGLDHAENI
LSYLDKSLCAELVCKEWRVTSQMLKLIETISNWRGSRHSLQIRHSETSKGVCLQYD
NKPPDGNAPNSFYRALYPKIIODIETISNWRGSRHSLQIRHSETSKGVCLQYD
DQKIVSLGRDNTIKIMDKNTLECKRIILTGHTSVLCLOYDERVILITGSSDSTVRVMDV
NTGEMVLTILHHCVAVLHFRNFMWVTCSKDRSIAVMDWASPTDITLRLVLGHRAA
VNVDFDKYIVSASGDTIKVWNTSCFVRLNCHKRGICLQYRDLVYSGSDN
TIRLWDIEGACLRLVGBHELVRCIRPDNKRIVSGAYDKIKVWDLVLAALDFRAPAG
TLCURTLVHSGRFRLOQDEFQIVSSSHDDTILINDFLNDPAAQAEPSPRSRTVTV
ISR"

BASE COUNT 628 a 467 c 513 g 543 t

CDS

ORIGIN

Query Match 100.0%; Score 2151; DB 6; Length 2151;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCGGTGGCTGGCGCTGGCCACCAAGGGCGCGCCGGGAGCGGACCCAGTGGCC 60
DB 1 TCGGTGGCTGGCGCTGGCCACCAAGGGCGCGCCGGGAGCGGACCCAGTGGCC 60
QY 61 TCGGGGATATGCGACCCCGCGAGCGGTGTGCAAGAGAGGACCTCAAGTTATGAAT 120
DB 61 TCGGGGATATGCGACCCCGCGAGCGGTGTGCAAGAGAGGACCTCAAGTTATGAAT 120
QY 121 TCCTCAGAGAGAGAGACTGTAATAATGGCGAACCCCTAGGAAGATAATACAGAGAAG 180
DB 121 TCCTCAGAGAGAGAGACTGTAATAATGGCGAACCCCTAGGAAGATAATACAGAGAAG 180
QY 181 AATTCATTAGACAGACATACAACAGCTGTGCCAGACTCTGCTPAAACCAAGAAAACAGTA 240
DB 181 AATTCATTAGACAGACATACAACAGCTGTGCCAGACTCTGCTPAAACCAAGAAAACAGTA 240
QY 241 TGTTTAGCAAGCACTGCTATCAAGACTGAGAAATGTGTGGCCAAAACAAAACCTTGCCAAT 300
DB 241 TGTTTAGCAAGCACTGCTATCAAGACTGAGAAATGTGTGGCCAAAACAAAACCTTGCCAAT 300
QY 301 GGCATTCAGATATGTTGCCCAAGCAACGAAACTCTCAGCAAGCTATGAAAAGGAA 360
DB 301 GGCATTCAGATATGTTGCCCAAGCAACGAAACTCTCAGCAAGCTATGAAAAGGAA 360
QY 361 AAGGAACGTGTGTCAATATCTTTGACGAGTGTGTCAGAGTCAGATCAAGTGGAAATTTGTG 420
DB 361 AAGGAACGTGTGTCAATATCTTTGACGAGTGTGTCAGAGTCAGATCAAGTGGAAATTTGTG 420
QY 421 GAACATCTTATATCCAAATGTGCTATTACCAACATGGGCACATAAACTCGTATCTTAAA 480
DB 421 GAACATCTTATATCCAAATGTGCTATTACCAACATGGGCACATAAACTCGTATCTTAAA 480
QY 481 CCTATGTTGCAGAGAGATTTTCAATCTGCTGACGCTCGGAGTTGGATCATATCGCT 540
DB 481 CCTATGTTGCAGAGAGATTTTCAATCTGCTGACGCTCGGAGTTGGATCATATCGCT 540
QY 541 GAGAACATCTGCTATACCTGGATGCCAAATCACTATGTGCTGCTGAACCTTGTGTGCAAG 600
DB 541 GAGAACATCTGCTATACCTGGATGCCAAATCACTATGTGCTGCTGAACCTTGTGTGCAAG 600

QY 601 GAATGGTACCAGTGAGCTCTGATGGCATCTCTGGAAGAGCTTATCAGAGAATGGTC 660
DB 601 GAATGGTACCAGTGAGCTCTGATGGCATCTCTGGAAGAGCTTATCAGAGAATGGTC 660
QY 661 AGGACAGATTCTCTGTGGAGAGGCTGGCAGAACGAGAGATGGGAGCAGTATTTATTC 720
DB 661 AGGACAGATTCTCTGTGGAGAGGCTGGCAGAACGAGAGATGGGAGCAGTATTTATTC 720
QY 721 AAAAACAACCTCTCTGAGGGGAATGCTCTCCCAACTCTCTTTTATAGACACTTTATCT 780
DB 721 AAAAACAACCTCTCTGAGGGGAATGCTCTCCCAACTCTCTTTTATAGACACTTTATCT 780
QY 781 AAAATTATACAAGACATTGAGACAATAGAAATCTAAATGGAGATGTGGAGACATAGTTTA 840
DB 781 AAAATTATACAAGACATTGAGACAATAGAAATCTAAATGGAGATGTGGAGACATAGTTTA 840
QY 841 CAGAAATTCATCTGCCGAAGTGAAACAAAGAAAGAGATTACCTGTTTACAGTATGAT 900
DB 841 CAGAAATTCATCTGCCGAAGTGAAACAAAGAAAGAGATTACCTGTTTACAGTATGAT 900
QY 901 CAGAAATAGTAAGCGGCTTCGAGACAACACAATCAAGATCTGGGATAAACAACATG 960
DB 901 CAGAAATAGTAAGCGGCTTCGAGACAACACAATCAAGATCTGGGATAAACAACATG 960
QY 961 GAATGCAAGGAAATTCATCAGGCCATACAGGTTCACTCTCTCTCCAGTATGATGAG 1020
DB 961 GAATGCAAGGAAATTCATCAGGCCATACAGGTTCACTCTCTCTCCAGTATGATGAG 1020
QY 1021 AGAGTGATCATACAGGATCATCGGATTCACCGGTGAGAGTGGGATCTAAATACAGT 1080
DB 1021 AGAGTGATCATACAGGATCATCGGATTCACCGGTGAGAGTGGGATCTAAATACAGT 1080
QY 1081 GAAATGCTAAACACGTTGATTCAACCAATGTGTAAGCAGTTCTGCACTTGGTTCAATAAT 1140
DB 1081 GAAATGCTAAACACGTTGATTCAACCAATGTGTAAGCAGTTCTGCACTTGGTTCAATAAT 1140
QY 1141 GGCATGATGTGACCTGCTCCAAAGATCGTTCAATTGTAAGCAGTTCTGCACTTGGTTCAATAAT 1200
DB 1141 GGCATGATGTGACCTGCTCCAAAGATCGTTCAATTGTAAGCAGTTCTGCACTTGGTTCAATAAT 1200
QY 1201 ACTCAGATTACCTCTCGAGGGTGTGTCGGAACCCGAGCTGCTCAATGTTGTAGAC 1260
DB 1201 ACTCAGATTACCTCTCGAGGGTGTGTCGGAACCCGAGCTGCTCAATGTTGTAGAC 1260
QY 1261 TTTGATGACAAAGTACATTTGTTCTGCACTCTGGGATAGAACTATAAAGGTATGGAACACA 1320
DB 1261 TTTGATGACAAAGTACATTTGTTCTGCACTCTGGGATAGAACTATAAAGGTATGGAACACA 1320
QY 1321 AGTACTTGTGAATTTGTAAGGACCTTAAATGGACACAAACGAGGCAATGCGCTGTTGCGAG 1380
DB 1321 AGTACTTGTGAATTTGTAAGGACCTTAAATGGACACAAACGAGGCAATGCGCTGTTGCGAG 1380
QY 1381 TACAGGACAGGCTGTGATGAGTGTCTCATCTGACAACTATCAGATATGGAACATA 1440
DB 1381 TACAGGACAGGCTGTGATGAGTGTCTCATCTGACAACTATCAGATATGGAACATA 1440
QY 1441 GAATGTGTCATGTTTACGAGTGTAGAGGCAATGAGGAAATGGTGTGTTGTAATCGA 1500
DB 1441 GAATGTGTCATGTTTACGAGTGTAGAGGCAATGAGGAAATGGTGTGTTGTAATCGA 1500
QY 1501 TTTGATTAACAGAGGATGATGATGAGGCTATGATGGAATAAATTAAGTGTGGGATCTTT 1560
DB 1501 TTTGATTAACAGAGGATGATGATGAGGCTATGATGGAATAAATTAAGTGTGGGATCTTT 1560
QY 1561 GTGGCTGTTTGGACCCCGTCTCTGAGGAGACTCTGCTCTACGAGCCCTGTGGAG 1620
DB 1561 GTGGCTGTTTGGACCCCGTCTCTGAGGAGACTCTGCTCTACGAGCCCTGTGGAG 1620
QY 1621 CATTCGGGAAGAGTTTTTTCGACTACAGTTTGTGAATTCAGATTTCTCAGTAGTTTCAAT 1680
DB 1621 CATTCGGGAAGAGTTTTTTCGACTACAGTTTGTGAATTCAGATTTCTCAGTAGTTTCAAT 1680

Db	541	GAGAACATTCTGTCTATACCTGGATGCCAAATCACTATGTGCTGCTGAACTTTGTGTGCAAG	600
QY	601	GAATGGTACCGAGTGAACTCTGATGGCATGCTGTGGAAGAACTTATCGAGAGAAATGGTC	660
Db	601	GAATGGTACCGAGTGAACTCTGATGGCATGCTGTGGAAGAACTTATCGAGAGAAATGGTC	660
QY	661	AGGACAGATTCTCTGTGGAGAGGCTTGGCAGAACGAGAGGATGGGACAGTATTTATTC	720
Db	661	AGGACAGATTCTCTGTGGAGAGGCTTGGCAGAACGAGAGGATGGGACAGTATTTATTC	720
QY	721	AAAAAACAACCTCTGACGGGAATGCTCTCCCAACTCTTTTATAGAGCACTTTATCCT	780
Db	721	AAAAAACAACCTCTGACGGGAATGCTCTCCCAACTCTTTTATAGAGCACTTTATCCT	780
QY	781	AAATATTATACAGACATTGAGACAAATAGAACTTAATTTGGAGATGTGGAAGACATAGTTTA	840
Db	781	AAATATTATACAGACATTGAGACAAATAGAACTTAATTTGGAGATGTGGAAGACATAGTTTA	840
QY	841	CAGAGAAATTCACCTCCGAAGTGAACAGCAAGCAAGGAGTTTACTGTTTACAGTATGATGAT	900
Db	841	CAGAGAAATTCACCTCCGAAGTGAACAGCAAGCAAGGAGTTTACTGTTTACAGTATGATGAT	900
QY	901	CAGAAATAGTAAGCGGCTTCGAGACAAACAACTCAAGATCTGGGATAAAACACATTG	960
Db	901	CAGAAATAGTAAGCGGCTTCGAGACAAACAACTCAAGATCTGGGATAAAACACATTG	960
QY	961	GAATGCAAGCAATTCACAGGCCATACAGGTTCACTCTCTCTCCAGTATGATGAG	1020
Db	961	GAATGCAAGCAATTCACAGGCCATACAGGTTCACTCTCTCTCCAGTATGATGAG	1020
QY	1021	AGAGTGATCAACAGGATCATCGGATTCACAGGTCAGAGTGTGGAGTGAATATACAGGT	1080
Db	1021	AGAGTGATCAACAGGATCATCGGATTCACAGGTCAGAGTGTGGAGTGAATATACAGGT	1080
QY	1081	GAATGCTAAACAGCTTGAATCACCATTGTGAAGCAGTTCTGACATTCGCTTTCAATTAAT	1140
Db	1081	GAATGCTAAACAGCTTGAATCACCATTGTGAAGCAGTTCTGACATTCGCTTTCAATTAAT	1140
QY	1141	GGCATGATGGTACCTGCTCAAGAGATCGTTCCATGCTGTATGGATATGGCTCCCA	1200
Db	1141	GGCATGATGGTACCTGCTCAAGAGATCGTTCCATGCTGTATGGATATGGCTCCCA	1200
QY	1201	ACTGACATTACCTCCGAGGCTGCTGTCGACACCGAGCTGCTGCAATGTTGTAGAC	1260
Db	1201	ACTGACATTACCTCCGAGGCTGCTGTCGACACCGAGCTGCTGCAATGTTGTAGAC	1260
QY	1261	TTTGATGACAGTACATTTGTTCTGCACTCTGGGATAGAACTATAAAGGTATGGAACACA	1320
Db	1261	TTTGATGACAGTACATTTGTTCTGCACTCTGGGATAGAACTATAAAGGTATGGAACACA	1320
QY	1321	AGTACTTGTAATTTGTAAGGACCTTAAATGGAACACAAAGGAGGATGCTGTTGCGAG	1380
Db	1321	AGTACTTGTAATTTGTAAGGACCTTAAATGGAACACAAAGGAGGATGCTGTTGCGAG	1380
QY	1381	TACAGGACAGGCTGTAGTGTGAGTGTGCTATCGAACACATATCAGATTTATGGACATA	1440
Db	1381	TACAGGACAGGCTGTAGTGTGAGTGTGCTATCGAACACATATCAGATTTATGGACATA	1440
QY	1441	GAATGTGTGCTGTTTACAGAGTGTAGAGGCCATGAGGAATTTGTTGCGTGTATTGGA	1500
Db	1441	GAATGTGTGCTGTTTACAGAGTGTAGAGGCCATGAGGAATTTGTTGCGTGTATTGGA	1500
QY	1501	TTTGATCAACAGAGATAGTCACTGGGGCTTATGATGGAAAAATTAAGGTGTGGATCTT	1560
Db	1501	TTTGATCAACAGAGATAGTCACTGGGGCTTATGATGGAAAAATTAAGGTGTGGATCTT	1560
QY	1561	GTGGCTGCTTTGGACCCCGTCTCTCGAGGACACTCTGTCTACGACCCCTTGTGGAG	1620
Db	1561	GTGGCTGCTTTGGACCCCGTCTCTCGAGGACACTCTGTCTACGACCCCTTGTGGAG	1620
QY	1621	CATTCCGGAAGATTTTTCAGTACATGATTTGATGAAATTCAGATTTGTCAGTATTCACAT	1680
Db	1621	CATTCCGGAAGATTTTTCAGTACATGATTTGATGAAATTCAGATTTGTCAGTATTCACAT	1680
Db	1681	GATGACAAATCTCTCATCTGGGACTTCTTAATGATCCAGCTGCCCAAGCTGAACCCCCC	1740
Db	1681	GATGACAAATCTCTCATCTGGGACTTCTTAATGATCCAGCTGCCCAAGCTGAACCCCCC	1740
QY	1741	CGTTCCCTTCTCGAACAATACACCTTCCAGATAAATAAACCATACACTGACCTCAT	1800
Db	1741	CGTTCCCTTCTCGAACAATACACCTTCCAGATAAATAAACCATACACTGACCTCAT	1800
QY	1801	ACTTCCCGCAGGACCCATTAAAGTTCCGGTATTTAACTATCTGCCAATACCAGGATGAGC	1860
Db	1801	ACTTCCCGCAGGACCCATTAAAGTTCCGGTATTTAACTATCTGCCAATACCAGGATGAGC	1860
QY	1861	AACAACAGTAAACAATCAAACTACTCCCGAGTTTCCCTGGACTAGCCGAGGACGAGGCTT	1920
Db	1861	AACAACAGTAAACAATCAAACTACTCCCGAGTTTCCCTGGACTAGCCGAGGACGAGGCTT	1920
QY	1921	TCGAGACTCTGTGTTGGACACAGTTCGTGAGTCGGCCCGAGGACGGTCTACTCAGCACA	1980
Db	1921	TCGAGACTCTGTGTTGGACACAGTTCGTGAGTCGGCCCGAGGACGGTCTACTCAGCACA	1980
QY	1981	ACTGACTCTTCAGTGTCTATCAGAAAGATGTCTTCTATCAATTTGTAATGATTTGGAAC	2040
Db	1981	ACTGACTCTTCAGTGTCTATCAGAAAGATGTCTTCTATCAATTTGTAATGATTTGGAAC	2040
QY	2041	TTTTAAACCTCCCT	2100
Db	2041	TTTTAAACCTCCCT	2100
QY	2101	CAGACAAAGGTGACTTAAATATATTTAGTGTGTTTGGCAGAAAAAATAA 2151	
Db	2101	CAGACAAAGGTGACTTAAATATATTTAGTGTGTTTGGCAGAAAAAATAA 2151	
RESULT 4			
AF129530			
LOCUS	AF129530	2151 bp	mRNA linear PRI 31-OCT-1999
DEFINITION	Homo sapiens chromosome 10 F-box protein FBW1A (FBW1A) mRNA, complete cds.		
ACCESSION	AF129530		
VERSION	AF129530.1	GI:6164609	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 2151) Cenciarelli, C., Chiau, D.S., Guardavaccaro, D., Parks, W., Vidal, M. and Pagano, M.		
TITLE	Identification of a family of human F-box proteins		
JOURNAL	Mol. Cell. Biol. 19 (20), 1177-1179 (1999)		
MEDLINE	10531035		
PUBMED	2 (bases 1 to 2151)		
REFERENCE	Chiau, D.S. and Pagano, M.		
AUTHORS	Direct Submission		
TITLE	Submitted (19-FEB-1999) Pathology, NYU Medical Center, 550 First Ave. MSB 548, New York, NY 10016, USA		
JOURNAL	Location/Qualifiers		
FEATURES	source		
source	1..2151		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/chromosome="10"		
	/map="10q24"		
	1..2151		
	/gene="FBW1A"		
	70..1779		
	/gene="FBW1A"		
	/function="regulates the stability of IkappaB alpha- and beta-catenin"		
	/note="contains one F-box domain and seven WD40 domains; forms an SCF with Skp1 and Cull1"		
gene			
CDS			

/codon_start=1									
/product="F-box protein Fbw1A"									
/protein_id="AA04464.1"									
/db_xref="GI:6164610"									
/translation="MDPAEAVLQEKALKFMNSSEDCNNGPPRKIIPEKNSLRQTV									
NSCARLCLNQTVCLASTAMKNTENCVAKTLANGTSMIVPKRKLSASGEKEKELCV									
KYFOWSESDOVEFVHLISQMYCHYOHGHINSYLKPMLODFITLALPARGLDHIENIK									
LSYLDASLCAELVCKEYWRVSDGMWLKXLERWVRDLSLRGLAERGGWGVLFK									
NKPDGAPNPSFTRALYPKLIQDIETIESNRCGRHSQRIHCRSETSKGVICYLQYD									
DOKIVSGLRNTIKIKWDKNTLECKRILITGHTSVLCLOYDERVITIGSSDSVTRVWDV									
NTGMLNLTILHICAVLHLRFNNGMVTCSKDRSIAWDMASPTDITLRLVLVGHRAA									
VVVWDFDQITVSASGDRITIKVWNTSTCFEVRILNGHKGRIACLOWDLVAVLGGSDN									
TIRLWDDTBCACILRVGLHEELVRCIRFDNKRIVSGAYDKIKWDLVAALDPRAPAG									
TLCLRLTIVHSGRVFRQLQFDEFQIVSSSHDDTILINDFLNDPAAQAEPRPSRITYTY									
ISR"									
BASE COUNT 628 a 467 c 513 g 543 t									
ORIGIN									
Query Match 100.0%; Score 2151; DB 9; Length 2151;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	TGCGTGGCTGGCGCTGGCACCAAGGGGGCGCCCGCGGAGAGCGGACCGAGTGGCC	60						
DB	1	TGCGTGGCTGGCGCTGGCACCAAGGGGGCGCCCGCGGAGAGCGGACCGAGTGGCC	60						
QY	61	TCGGCGATTATGGACCCCGCGGAGCGGTGTGCAAGAGAGAGGCACTCAAGTTTATGAAT	120						
DB	61	TCGGCGATTATGGACCCCGCGGAGCGGTGTGCAAGAGAGAGGCACTCAAGTTTATGAAT	120						
QY	121	TCCTCAGAGAGAGAGAGCTGTAATAATGGGGAACCCCTAGGAAGTAATACAGAGAAG	180						
DB	121	TCCTCAGAGAGAGAGAGCTGTAATAATGGGGAACCCCTAGGAAGTAATACAGAGAAG	180						
QY	181	AATTCACCTTAGACAGACATACACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTA	240						
DB	181	AATTCACCTTAGACAGACATACACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTA	240						
QY	241	TGTTTAGCAAGCACTGTATGAAGCTGAGATTTGTGTGCCAAACAAACAAACCTTGCCAT	300						
DB	241	TGTTTAGCAAGCACTGTATGAAGCTGAGATTTGTGTGCCAAACAAACAAACCTTGCCAT	300						
QY	301	GGCACTTCAGTATGATTGTGCCCAAGCAACGGAAACTCTCAGCAAGCTATGAAAGGAA	360						
DB	301	GGCACTTCAGTATGATTGTGCCCAAGCAACGGAAACTCTCAGCAAGCTATGAAAGGAA	360						
QY	361	AAGGAACCTGTGTCAATACTTTGAGCAGTGGTCAGAGTCAGATCAAGTGGAAATTGTG	420						
DB	361	AAGGAACCTGTGTCAATACTTTGAGCAGTGGTCAGAGTCAGATCAAGTGGAAATTGTG	420						
QY	421	GAACATCTTATATCCAAATGTCTCAATACCAATGGGCACATAAACTCGTATCTTAA	480						
DB	421	GAACATCTTATATCCAAATGTCTCAATACCAATGGGCACATAAACTCGTATCTTAA	480						
QY	481	CCTATGTTGAGAGAGATTTCAATACTCTGCGAGTGGTCAGAGTCAGATCAATTCGCT	540						
DB	481	CCTATGTTGAGAGAGATTTCAATACTCTGCGAGTGGTCAGAGTCAGATCAATTCGCT	540						
QY	541	GAGAACATCTGTCTATACCTGGATGCCAAATCACTATGTCTGCTGAATCTGTGTGCAAG	600						
DB	541	GAGAACATCTGTCTATACCTGGATGCCAAATCACTATGTCTGCTGAATCTGTGTGCAAG	600						
QY	601	GAATGGTACGAGTGAACCTCTGATGGCAATGCTGTGGGAAGAGCTTATTCGAGAGAATGGTC	660						
DB	601	GAATGGTACGAGTGAACCTCTGATGGCAATGCTGTGGGAAGAGCTTATTCGAGAGAATGGTC	660						
QY	661	AGGACAGATTTCTCTGTGAGAGGCTGTGAGAAAGAGAGATGGGACAGATTTATTTC	720						
DB	661	AGGACAGATTTCTCTGTGAGAGGCTGTGAGAAAGAGAGATGGGACAGATTTATTTC	720						
QY	721	AAAAACAAACCTCTGACGGGAATGCTCTCCCAACTCTTTTATAGACACCTTTATCT	780						
DB	721	AAAAACAAACCTCTGACGGGAATGCTCTCCCAACTCTTTTATAGACACCTTTATCT	780						

QY	781	AAAAATATTAAGACATTCGAGCAATAGAACTTAATGGAGATGTGGAAGACATAGTTTA	840						
DB	781	AAAAATATTAAGACATTCGAGCAATAGAACTTAATGGAGATGTGGAAGACATAGTTTA	840						
QY	841	CAGAGAAATCACTGGCGGAGTCAAAACAGCAAGAGGTTTACTGTTTACAGTATGATGAT	900						
DB	841	CAGAGAAATCACTGGCGGAGTCAAAACAGCAAGAGGTTTACTGTTTACAGTATGATGAT	900						
QY	901	CAGAAATATAGTAAGCGGCTTCAGACAAACAACAATCAAGATCTGGGATAAAAACACATTTG	960						
DB	901	CAGAAATATAGTAAGCGGCTTCAGACAAACAACAATCAAGATCTGGGATAAAAACACATTTG	960						
QY	961	GAATGCAAGGGAATCTCAAGGCGCATACAGTTTCAGTCTCTCTCTCCAGTATGATGAG	1020						
DB	961	GAATGCAAGGGAATCTCAAGGCGCATACAGTTTCAGTCTCTCTCTCCAGTATGATGAG	1020						
QY	1021	AGAGTGATCATTAACAGGATCATCGGATTCACGGTCAGAGTGTGGGATGTAATACAGGT	1080						
DB	1021	AGAGTGATCATTAACAGGATCATCGGATTCACGGTCAGAGTGTGGGATGTAATACAGGT	1080						
QY	1081	GRAATCTAAACACAGTTGATTCACCAATTTGGAAGCAGTTCTGCACCTTCGGTTTCAATAAT	1140						
DB	1081	GRAATCTAAACACAGTTGATTCACCAATTTGGAAGCAGTTCTGCACCTTCGGTTTCAATAAT	1140						
QY	1141	GGCATGATGGTGACCTGTCTCCAAAGATCGTTCCATTGCTATGGGATATGCGCTCCCCA	1200						
DB	1141	GGCATGATGGTGACCTGTCTCCAAAGATCGTTCCATTGCTATGGGATATGCGCTCCCCA	1200						
QY	1201	ACTGACATTAACCTTCGGAGGGTGTGTGCGACACCGAGCTGTCTCAATGTTGTAGAC	1260						
DB	1201	ACTGACATTAACCTTCGGAGGGTGTGTGCGACACCGAGCTGTCTCAATGTTGTAGAC	1260						
QY	1261	TTTGTATGACAAAGTACATTTGCTCATCTGGGATGAACTATAAAGGTATGGAACACA	1320						
DB	1261	TTTGTATGACAAAGTACATTTGCTCATCTGGGATGAACTATAAAGGTATGGAACACA	1320						
QY	1321	AGTACTGTGAAATTTGTAAGGACCTTAATGAGACAAACAGGCAITGCTGTTTGCAG	1380						
DB	1321	AGTACTGTGAAATTTGTAAGGACCTTAATGAGACAAACAGGCAITGCTGTTTGCAG	1380						
QY	1381	TACAGGACAGGCTGTGTAGTGGCTCATCTGACAAACACTATCAGATATGGGACATA	1440						
DB	1381	TACAGGACAGGCTGTGTAGTGGCTCATCTGACAAACACTATCAGATATGGGACATA	1440						
QY	1441	GAATGTTGTCATGTTTACAGTGTAGAGCCCATGAGAAATTTGGTGGCTGTTATTCGA	1500						
DB	1441	GAATGTTGTCATGTTTACAGTGTAGAGCCCATGAGAAATTTGGTGGCTGTTATTCGA	1500						
QY	1501	TTTGATAACAAGAGATAGTCAGTGGGCTATGATGGAATAATTAAGTGTGGGATCTT	1560						
DB	1501	TTTGATAACAAGAGATAGTCAGTGGGCTATGATGGAATAATTAAGTGTGGGATCTT	1560						
QY	1561	GTGGCTGCTTTGGACCCCGTCTCTGAGGAGACATCTCTGTCTACGGAACCTTGTGGAG	1620						
DB	1561	GTGGCTGCTTTGGACCCCGTCTCTGAGGAGACATCTCTGTCTACGGAACCTTGTGGAG	1620						
QY	1621	CATTCCGGGAAGAGTTTTCGACTACAGTTTGTGATGAATCCAGATTGTTCAGTATTCAT	1680						
DB	1621	CATTCCGGGAAGAGTTTTCGACTACAGTTTGTGATGAATCCAGATTGTTCAGTATTCAT	1680						
QY	1681	GATGACAAATCTCTGGAACATACACCTTACATCTCCAGATAAATCACTACCTGACCTCAT	1740						
DB	1681	GATGACAAATCTCTGGAACATACACCTTACATCTCCAGATAAATCACTACCTGACCTCAT	1740						
QY	1741	CGTTCCCTTCTCGAACATACACCTTACATCTCCAGATAAATCACTACCTGACCTCAT	1800						
DB	1741	CGTTCCCTTCTCGAACATACACCTTACATCTCCAGATAAATCACTACCTGACCTCAT	1800						
QY	1801	ACTTGCCCGAGGACCCATTAAGTTGCGGTATTTAACTGATCTGCCAATACCGAGATGAGC	1860						
DB	1801	ACTTGCCCGAGGACCCATTAAGTTGCGGTATTTAACTGATCTGCCAATACCGAGATGAGC	1860						

	TGALTAFQNSBEREDCNNGPPRKIIPEKNSLRQTYNSCARLCLNQBTVCVLASTPMKTI ENCVAKTLANGTSMIVPKQRKSAYEKEKLCKVYFEQWSBDSQDVEFVHLISQM CHYOQHINSLYLKMLQDRDFITALPARGDLHIAENILSYLDXKSLCAELVCKEWVRV TSDGLNWKRLIERMVRTLSLWRGAEREGWGYLPFKNPDPDNAPPNSFYRALYPKII ODIETLESNNRGESLORIHCRSETSKGVLCLOYDDOKIVSGLRDNTIKWDKNLTLE CKRILIGHTSVLCLQDYDERVITITSSSSTVRVDVTGEMTLNIHHCEAVLHLREN NMGMVTSKORSIAVMDWASPIDITLRRVLVGHRAAENVVDFFDKYIVSASGGORTKV WNMTSEFVRLNHKGHTACIQVDRDLRVSGSDNTIRLWDIECGACLVLBGHEEL VICIFPDNRKIVSGAYDGKIKWDLVAALDPRAPAGTLCLRFLVHSGRVRFLOQDEF QIVSSSHDDTILIIMFLNDPAQAEPSPSRRTVTVISR"	BASE COUNT 719 a	493 c	530 g	566 t	ORIGIN
	Query Match 94.4%; Score 2030.4; DB 9; Length 2308; Best Local Similarity 99.7%; Pred. No. 0; Matches 2034; Conservative 0; Mismatches 6; Indels 0; Gaps 0;					
Qy	112 TTATCAAATTCCTCAGAGAGAAGACTGTAAATAATGGCAACCCCTAGGAAGATAATA 171					
Db	188 TTTCCAGAAATTCCTCAGAGAGAAGACTGTAAATAATGGCAACCCCTAGGAAGATAATA 247					
Qy	172 CCAGAGAGAAATTCATTAGACAGACATACAACAGCTGTGCAGACTCTGCTTAAACCAA 231					
Db	248 CCAGAGAGAAATTCATTAGACAGACATACAACAGCTGTGCAGACTCTGCTTAAACCAA 307					
Qy	232 GAACAGTAGTTTACCAAGCACTGCTATGAAGACTGAGAAATGTGTGCCCAACAAA 291					
Db	308 GAACAGTAGTTTACCAAGCACTGCTATGAAGACTGAGAAATGTGTGCCCAACAAA 367					
Qy	292 CTTGCCAAATGGCACTTCCAGTAGATGATGTGCCAACAGCAACGGAAACTCTCAGCAAGCTAT 351					
Db	368 CTTGCCAAATGGCACTTCCAGTAGATGATGTGCCAACAGCAACGGAAACTCTCAGCAAGCTAT 427					
Qy	352 GAACAGAAAGCAACTGTGTGTCARAATCTTTTGACGAGTGTGCAGAGTCAGATCAAGTG 411					
Db	428 GAACAGAAAGCAACTGTGTGTCARAATCTTTTGACGAGTGTGCAGAGTCAGATCAAGTG 487					
Qy	412 GAATTTGTGGAACATCTTATATCCCAAATGTGTCAATTACCACATGGGCACATAAACTCG 471					
Db	488 GAATTTGTGGAACATCTTATATCCCAAATGTGTCAATTACCACATGGGCACATAAACTCG 547					
Qy	472 TATCTTAACCTATGTGCAGAGAGATTTTCAATCTGCTCTCCGAGCTCGGGATGSGAT 531					
Db	548 TATCTTAACCTATGTGCAGAGAGATTTTCAATCTGCTCTCCGAGCTCGGGATGSGAT 607					
Qy	532 CATATGCTGAGAACATTTCTGTATACCTGGATGCCAAATCACTATGTGCTGCTGAACTT 591					
Db	608 CATATGCTGAGAACATTTCTGTATACCTGGATGCCAAATCACTATGTGCTGCTGAACTT 667					
Qy	592 GTGTGCAGAGATGTGACGAGTACCTCTGTATGGCATGCTGTGAAAGAGCTTATCGAG 651					
Db	668 GTGTGCAGAGATGTGACGAGTACCTCTGTATGGCATGCTGTGAAAGAGCTTATCGAG 727					
Qy	652 AGAATGCTCAGGACAGATTTCTGTGAGAGGCTGTGGCAGAACGAGAGGATGGGACAG 711					
Db	728 AGAATGCTCAGGACAGATTTCTGTGAGAGGCTGTGGCAGAACGAGAGGATGGGACAG 787					
Qy	712 TATTTATTTAAAAACAAAACCTCTGTACGGGAATGTCTCTCCAACTCTTTTTATAGACA 771					
Db	788 TATTTATTTAAAAACAAAACCTCTGTACGGGAATGTCTCTCCAACTCTTTTTATAGACA 847					
Qy	772 CTTTATCTCTAAAATTTATACAGACATTTGACAAATAGATCTAAATGGAGATGTGGAAGA 831					
Db	848 CTTTATCTCTAAAATTTATACAGACATTTGACAAATAGATCTAAATGGAGATGTGGAAGA 907					
Qy	832 CATAGTTTACAGAGAAATTCATCTGCCGAAGTGAACCAAGCAAAAGGAGTTTACTGTTTACAG 891					
Db	908 CATAGTTTACAGAGAAATTCATCTGCCGAAGTGAACCAAGCAAAAGGAGTTTACTGTTTACAG 967					
Qy	892 TATGATGATCAGAAAAATAGTAAGCGCTTCGAGACACACAAATCAGATCTGGGATAAA 951					
Db	968 TATGATGATCAGAAAAATAGTAAGCGCTTCGAGACACACAAATCAGATCTGGGATAAA 1027					

Db	1222	ACTGACATCACCCCTCAGGAGGGTCTCTGGTGGGACACCGAGCTCGGCTCAATGTTGTAGAC	1281
Qy	1261	TTTGATGACAAAGTACATGTTTCTGCACTCTGGGATAGAACTATAAGGTATGGAACACA	1320
Db	1282	TTTGATGACAAAGTACATGTTTCTGCACTCTGGGATAGAACTATAAGGTATGGAACACA	1341
Qy	1321	AGTACTTGTGAATTTGTAAGACCTTAAATGGACACAAACGAGGCAATGCTCTTTGGAG	1380
Db	1342	AGTACTTGTGAATTTGTAAGACCTTAAATGGGACAAAGGCTGGCATCGCTCTTTGGAG	1401
Qy	1381	TACAGGACAGGCTGTGTAGTGCAGTGGCTCATCTGCACACACACTATCAGATTATGGGACATA	1440
Db	1402	TACAGACAGGCTGTGTAGTGCAGGCTCTCTGCACACACACTATCAGCTGTGGGACATA	1461
Qy	1441	GAATGTGTGCATGTTTACGAGTGTTAAGAGGCGCATGAGGAATGTCGTGTTATTCGA	1500
Db	1462	GAGTGTGGAGCATGCTCGAGTGTGTGAGGGCGCTATGAGGAGTGTGTGCTGCAATCGA	1521
Qy	1501	TTTGATACAGAGGATAGTACAGTGGGCGCTATGATGGAATAATTAAGTGTGGATCTT	1560
Db	1522	TTTGATACAAAGGATAGTACAGGAGCGCTATGATGGGAAATTAAGTGTGGATCTT	1581
Qy	1561	GTGGCTGTTTGGACCCCGCTCTCTCGAGGACACTCTGTCTACGACCTTGTGGAG	1620
Db	1582	ATGGCTGTTTGGACCCCGCTCTCTCGAGGAGCTCTCTGTCTGGGACACTTGTGGAG	1641
Qy	1621	CATTCCGGAAGATGTTTTCGACTACAGTTCATGTAATTCAGATTGTCACTAGTTCACAT	1680
Db	1642	CATTCTGGAAGATGTTTTCGCTCTCAGTTCATGTAATTCAGATTGTCACTAGTTCACAT	1701
Qy	1681	GATGACAAATCTCATCTGGGACTTCTAAATGATCCAGCTGCCAAGCTGAACCCGCC	1740
Db	1702	GATGACAAATCTCATCTGGGACTTCTGATGATCCAGCTGCTCAAGCTGAACCCGCC	1761
Qy	1741	CGTCCCTCTTCGAAACATACACCTACATCTCCAGATAAATAACATACACTACATCAT	1800
Db	1762	CGCTCCCTCTTCGAAACATACACCTACATCTCCAGATAAATAACATACACTACATCAT	1821
Qy	1801	ACTTGCCAGACCCCATTAAGTTCGCTATTTAACTGATCTGCTCAATACAGATACAGATGAGC	1860
Db	1822	AATTGCCAGGATTCGTTAATGTTGCGATATTTAAACAGCTGCCAAGACAGATGAAC	1881
Qy	1861	AACAACAGTAACTAACTACTGCCAGTTCCTCCCTG--GACTAGCCGAGGACGAGGC	1918
Db	1882	-----AACATCAACTCTTACCAGGATTCCTCCGACGATGACGAGGACGAGGC	1932
Qy	1919	TTTGAGACTCTGTGGACACAGTGTGTGCTGAGTGTGCGCCAGGACGCTCTACTAGCA	1978
Db	1933	TTTGAGACTCTGTGTGGACACAGTGTGCTGAGTGTGCGCCAGGACGCTCTACTAGCA	1990
Qy	1979	CAACTGACTGTGCTGCTATCAGAGATGCTCTTATCAATTTGGAATGATGGA	2038
Db	1991	--ACCGGTGCTGCTGCTATCAGAGATGCTCTTATCAATTTGGAATGATGGA	2047
Qy	2039	ACTTTTAAAGCTCCCTCCTCC--TCCTTTCACTCTGACCTAGTCTTTTCCCAT	2095
Db	2048	ACTTTTAAAGCTCCCTCCTCCCTTCCCTTCCCTTCCCTGACCTGCTTTCCCTCCCAT	2107
Qy	2096	-GGTTCACAGAAAGTCACTTATAATATTTAGTGTGTTTGGCAGAAAAAAA	2151
Db	2108	GGTTCACAGAAAGTCACTTATAATATTTAGTGTGTTTGGCAGAAAAAAA	2164
RESULT 8			
AF101784			
LOCUS AF101784 1818 bp mRNA linear PRI 21-JAN-1999			
DEFINITION Homo sapiens b-TRCP variant E3RS-IkappaB mRNA, partial cds.			
ACCESSION AF101784			
VERSION AF101784.1 GI:4165135			
KEYWORDS			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
1 (bases 1 to 1818)			
Yaron, A., Hatzubai, A., Davis, M., Lavon, I., Amit, S., Manning, A. M.,			
Andersen, J. S., Mann, M., Mercurio, F. and Ben-Neriah, Y.			
Identification of the receptor component of the			
IkappaBalpha-ubiquitin ligase			
Nature 396 (1998), 590-594 (1998)			
99075339			
9859996			
2 (bases 1 to 1818)			
Yaron, A., Hatzubai, A., Mercurio, F., Mann, A. M., Andersen, J. S.,			
Mann, K. and Ben-Neriah, Y.			
Direct Submission			
Submitted (27-OCT-1998) Immunology, Hebrew University-Hadassah			
Medical School, Jerusalem, Israel			
Location/Qualifiers			
1. 1818			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
<1. 1818			
/note="beta-transducin repeats-containing protein;			
receptor component of IkappaBalpha-ubiquitin ligase; contains			
F-box and WD regions"			
/codon_start=1			
/product="b-TRCP variant E3RS-IkappaB"			
/protein_id="AA008702.1"			
/db_xref="GI:4165136"			
/translation="MDFAEAVLQKALFKFMSRSLWLGCSLADSPSLRLCLVNP			
TGALTAFQNSEREDCNPPPKIIPKNSLQTYNSCARCLNGETVCLASTAMKT			
ENCVAKTKLANGTSSMIVKORSLSEYKEKELCKVKEFQMSQVEFVHLISAMT			
CHVGHGINSYLKPMQLQDITLALPARGLDHIAENILSYLDAKSLCAELVCKEWRV			
TSQMLWKLLIEMVSTDSLRGLAERGGGOVLKPKPDGNAPNSFVRLVYPKII			
QDLETIESWRCRHSRLQRCRSETSKGVYCLQYDDQKIVSLRLNTIKIWDKNTLE			
CKRLTGHVSILQIDERVITIGSSDSSTVRVMDVNTVMELNLLIHHCEAVLHLRFN			
NGMVTCKSDRSIAVMDASPTDITLRLVLRVHRAAVNVVDFDKYIVASGDRITKV			
WNTSTCFEVLNLRGKGIACQYRDLVVGSSDNTIRLWDIECCAGLRVLEHSEL			
VRCIRFQNKIRVSGVDGKIKVMDLVAALDPAPAGTLCRLTVEHSRGRVRLQPFDE			
QIVSSSHDDIILWDLFNDPAQAEPRSPRTYTVISR"			
BASE COUNT 532 a 385 c 445 g 456 t			
ORIGIN			
Query Match 77.3%; Score 1663.2; DB 9; Length 1818;			
Best Local Similarity 99.8%; Pred. No. 0;			
Matches 1665; Conservative 0; Mismatches 3; Indels 0; Gaps 0;			
Qy	112	TTTATGAATTCCTCAGAGAGAGAGACTGTAATAATGGCGAACCCCTAGGAGAGATAA	171
Db	151	TTCCAGAAATTCCTCAGAGAGAGAGACTGTAATAATGGCGAACCCCTAGGAGAGATAA	210
Qy	172	CCAGAGAGAAATTCACCTTAGACACATACACAGCTGTGCCAGACTCTGCTTTAAACCAA	231
Db	211	CCAGAGAGAAATTCACCTTAGACACATACACAGCTGTGCCAGACTCTGCTTTAAACCAA	270
Qy	232	GAACAGTATGTTTAGCAGACACTGCTATGAAAGTGTGAGTGTGTTGGCCAAACAAA	291
Db	271	GAACAGTATGTTTAGCAGACACTGCTATGAAAGTGTGAGTGTGTTGGCCAAACAAA	330
Qy	292	CTTGCAATGCACTTCCTCAGTATGATTGTGCCCAAGCAACGAAATCTCAGCAAGCTAT	351
Db	331	CTTGCAATGCACTTCCTCAGTATGATTGTGCCCAAGCAACGAAATCTCAGCAAGCTAT	390
Qy	352	GAAGAGGAAAGGAACTGTGTGTCAAATCTTTGAGCAGTGGTTCAGAGTTCAGATCAAGTG	411
Db	391	GAAGAGGAAAGGAACTGTGTGTCAAATCTTTGAGCAGTGGTTCAGAGTTCAGATCAAGTG	450
Qy	412	GAATTTGTGGAACATCTTATATCCAAATGTGTCAATACCAATGCGGCATATAACTCG	471
Db	451	GAATTTGTGGAACATCTTATATCCAAATGTGTCAATACCAATGCGGCATATAACTCG	510
Qy	472	TATCTTAAACCTATGTTGCGAGAGATTTCATAACTGCTCTGCCAGCTGGGATTTGGAT	531
Db	511	TATCTTAAACCTATGTTGCGAGAGATTTCATAACTGCTCTGCCAGCTGGGATTTGGAT	570

QY	532	CATATCGCTGAGAACATTTCTGTCTATACCTGGATGCGCAATCACTATGTGCTGCTGAACTT	591
Db	571	CATATCGCTGAGAACATTTCTGTCTATACCTGGATGCGCAATCACTATGTGCTGCTGAACTT	630
QY	592	GTGTGCAAGGAATGCTACCGAGTGACCTCTGATGCGCATGCTGTGGAAGACTTATCGAG	651
Db	631	GTGTGCAAGGAATGCTACCGAGTGACCTCTGATGCGCATGCTGTGGAAGACTTATCGAG	690
QY	652	AGAATGCTCAGGACAGATTTCTGTGTGAGAGCCCTGGCAGAACGAAGAGGATGGGACAG	711
Db	691	AGAATGCTCAGGACAGATTTCTGTGTGAGAGCCCTGGCAGAACGAAGAGGATGGGACAG	750
QY	712	TATTTATCAAAACAAACCTCTGACGGGAATGCTCTCCCACTCTTTTATGAGCA	771
Db	751	TATTTATCAAAACAAACCTCTGACGGGAATGCTCTCCCACTCTTTTATGAGCA	810
QY	772	CTTTATCTCTAAATATATCAAGACATTTGAGCAATAGAACTTAATTTGGAGATGTGGAAGA	831
Db	811	CTTTATCTCTAAATATATCAAGACATTTGAGCAATAGAACTTAATTTGGAGATGTGGAAGA	870
QY	832	CATAGTTTACAGAGATTTCACTCCCAAGTGAAACGAAGAGGATTTACTGTTTACAG	891
Db	871	CATAGTTTACAGAGATTTCACTCCCAAGTGAAACGAAGAGGATTTACTGTTTACAG	930
QY	892	TATGATGATCAGAAAAATAGTAAGCGGCTTTGAGACCAACCAATCAAGATCTGGGATAAA	951
Db	931	TATGATGATCAGAAAAATAGTAAGCGGCTTTGAGACCAACCAATCAAGATCTGGGATAAA	990
QY	952	AACACATGGAATGCAAGCGAATTTCTCAGAGCCATACAGTTGAGTCTCTGCTCCAG	1011
Db	991	AACACATGGAATGCAAGCGAATTTCTCAGAGCCATACAGTTGAGTCTCTGCTCCAG	1050
QY	1012	TATGATCAGAGAGTGATCATACAGGATCATCGGATTCACCGTTCAGAGTGTGGGATGA	1071
Db	1051	TATGATCAGAGAGTGATCATACAGGATCATCGGATTCACCGTTCAGAGTGTGGGATGA	1110
QY	1072	AATACAGGTGAATGCTTAACAGGTGATTCACCATGTTGAAGCATCTTCATCTGCGT	1131
Db	1111	AATACAGGTGAATGCTTAACAGGTGATTCACCATGTTGAAGCATCTTCATCTGCGT	1170
QY	1132	TTCAATTAATGGCATGATGTGACCTGCTCCAAAGATCGTTCCATTCGTTGATGGGATG	1191
Db	1171	TTCAATTAATGGCATGATGTGACCTGCTCCAAAGATCGTTCCATTCGTTGATGGGATG	1230
QY	1192	GCCTCCCAACTGACATTAACCTCCGAGGGTCTGCTCGACACCGAGCTGCTGCAAT	1251
Db	1231	GCCTCCCAACTGACATTAACCTCCGAGGGTCTGCTCGACACCGAGCTGCTGCAAT	1290
QY	1252	GTGTAGACTTTGATGACAGTACATTTGTTCTGTCATCTGGGGATAGAACTATAAGGTA	1311
Db	1291	GTGTAGACTTTGATGACAGTACATTTGTTCTGTCATCTGGGGATAGAACTATAAGGTA	1350
QY	1312	TGGAACACAGTACTTGTGAATTTGAAGCCTTAATGGAACACAAAGGAGGATTTGCC	1371
Db	1351	TGGAACACAGTACTTGTGAATTTGAAGCCTTAATGGAACACAAAGGAGGATTTGCC	1410
QY	1372	TGTTTTCAGTACAGGACAGGCTGGTAGTGAGTGCGTCTATCTGCAACACTATCAGATTA	1431
Db	1411	TGTTTTCAGTACAGGACAGGCTGGTAGTGAGTGCGTCTATCTGCAACACTATCAGATTA	1470
QY	1432	TGGGACATAGAAATGTTGGTGTGATTTTACAGTGTGTAGAGCCATGAGGAATTTGTCGCT	1491
Db	1471	TGGGACATAGAAATGTTGGTGTGATTTTACAGTGTGTAGAGCCATGAGGAATTTGTCGCT	1530
QY	1492	TGTATTCGATTTGATTAACAGAGGATAGTCAGTGGGCTTATGATGGAAAAATTAAGTG	1551
Db	1531	TGTATTCGATTTGATTAACAGAGGATAGTCAGTGGGCTTATGATGGAAAAATTAAGTG	1590
QY	1552	TGGGATCTTGTGCTGCTTTGGACCCCGTGTCTCTGCAAGGACACTCTGCTCTACGACC	1611
Db	1591	TGGGATCTTGTGCTGCTTTGGACCCCGTGTCTCTGCAAGGACACTCTGCTCTACGACC	1650
QY	1612	CTTTGGAGCAATTCGGAGAGTTTTTCGACTACATTTTGATGAATTCAGATTTGTCACT	1671
Db	1651	CTTTGGAGCAATTCGGAGAGTTTTTCGACTACATTTTGATGAATTCAGATTTGTCACT	1710
QY	1672	AGTTTCATGATGACACAATCCTCATCTGGAGCTTCTCTAAATGATCAGCTGCCAAGCT	1731
Db	1711	AGTTTCATGATGACACAATCCTCATCTGGAGCTTCTCTAAATGATCAGCTGCCAAGCT	1770
QY	1732	GAACCCCGGCTTCCCTTTCGAAACATACCATCATCTCCAGATAA	1779
Db	1771	GAACCCCGGCTTCCCTTTCGAAACATACCATCATCTCCAGATAA	1818
RESULT 9			
BC003989			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
2081 bp mRNA linear ROD 16-APR-2003			
Mus musculus beta-transducin repeat containing protein, mRNA (cdna			
clone MGC:7517 IMAGE:3491843), complete cds.			
BC003989			
GI:13278339			
MGC			
Mus musculus (house mouse)			
Mus musculus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
1 (bases 1 to 2081)			
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,			
Klausner, R.D., Collins, F.S., Wagner, L.H., Shenmen, C.M., Schuler, G.D.,			
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,			
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,			
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,			
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,			
Schreitz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,			
Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J.,			
Abrahamson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,			
Kernman, K.J., Mulek, J.A., Gunaratne, P.H., Richards, S.,			
Worley, K.C., Hale, S.S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,			
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,			
Faney, J.J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,			
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,			
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,			
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,			
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E.,			
Schneerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.			
Generation and initial analysis of more than 15,000 full-length			
human and mouse cDNA sequences			
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)			
22388257			
12477932			
2 (bases 1 to 2081)			
Strausberg, R.			
Direct Submission			
Submitted (28-FEB-2001) National Institutes of Health, Mammalian			
Gene Collection (MGC), Cancer Genomics Office, National Cancer			
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,			
USA			
NIH-MGC Project URL: http://mgc.nci.nih.gov			
Contact: MGC help desk			
Email: cgapbs-remail.nih.gov			
Tissue Procurement: Gilbert Smith, Ph.D.			
cDNA Library Preparation: Life Technologies, Inc.			
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)			
DNA Sequencing by: Sequencing Group at the Stanford Human Genome			
Center, Stanford University School of Medicine, Stanford, CA 94305			
Web site: http://www-ehgc.stanford.edu			
Contact: (Dickson, Mark) mdpaxil@stanford.edu			
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,			
R. M.			
Clone distribution: MGC clone distribution information can be found			
through the I.M.A.G.E. Consortium/LINL at: http://image.llnl.gov			
Series: IRAK Plate: 8 Row: 0 Column: 21			
This clone was selected for full length sequencing because it			

passed the following selection criteria: matched mRNA gi: 6753209.

FEATURES	source
Location/Qualifiers	
1..2081	
/organism="Mus musculus"	
/mol_type="mRNA"	
/strain="CZECH II"	
/db_xref="taxon:10090"	
/clone="MGC:7517 IMAGE:3491843"	
/tissue type="Mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMTV."	
/clone_lib="NCI CGAP_Lu29"	
/lab_host="DH10B"	
/note="Vector: pCMV-SPORT6"	
1..2081	
/gene="Btrc"	
/note="synonyms: PWD1, Slimb, b-TrCP, mSlimb"	
/db_xref="LocusID:12234"	
/db_xref="MGI:1338871"	
7..11716	
/codon_start=1	
/product="beta-transducin repeat containing protein"	
/protein_id="AAH03989.1"	
/db_xref="GI:13278340"	
/db_xref="LocusID:12234"	
(translation="MDPAEVLQEKALKFVNSSEEDCNNGEPKLIPEKNSLRQTY NSCARLCINQSTVCLTSTAMKTENCYAKAKLANGTSSMIVPKOKLSAYEKEKELCV KTFQWSSDQAEVFEHLISQMHYQGHSHLYLQKMLQDFITAPLAFGRGQHLAENI LSYDAKSLDAEIVCKEWRVYSDGMLWKLIERMVYTDLSMRGAERGGQVLYLFK NKPPENAPPSFVRLAYPKIIODIETIESNWRCHSLQRIHCRSETSKGVYCLQYD DOKI VSGLRDNTIKIMDKSTLEKRIILTGHTGSLVLCQYDERVITITGSSDSVTRVMDV NAGEMLTLIHCEAVLHLRFNNGMWTCSKORSIAVMDWASPTDITLRLVLVGHRAA VNVVDFDKYIVSASGDRTIKVMNTSCEFTUHLNGHKGIGACLOVDRDLVUSGSDN TIRLWDIEGCACLVLEGHBEIVRCIRFDNKRIVSGAYDKIKIWMALMDPRAPAG TUCRLTLVHSGRVRFLQDFQIVSSSHDDTILINDFLNDPAHAHPEPRSPRTITY I5R"	
BASE COUNT	568 a 502 c 529 g 482 t
ORIGIN	
Query Match	75.8%; Score 1630.4; DB 10; Length 2081;
Best Local Similarity	88.5%; Pred. No. 0;
Matches 1853; Conservative	0; Mismatches 221; Indels 20; Gaps 7;
QY	64 GCGATTATGGACCCGCGGAGCGGTGCTGCAAGAGAGGCACTCAAGTTTATGAATTC 123
DB	1 GCGATTATGGACCCGCGGAGCGGTGCTGCAAGAGAGGCACTCAAGTTTATGAATTC 60
QY	124 TCAGAGAGAGAGAGCTGTAATAATGCGCAACCCCTAGGAGAGATAATACAGAGAGAAT 183
DB	61 TCAGAGAGAGAGAGCTGTAATAATGCGCAACCCCTAGGAGAGATAATACAGAGAGAAT 120
QY	184 TCACCTTAGACAGACATACACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTATGT 243
DB	121 TCACCTTAGACAGACTTACACAGCTGTGCCAGCTTTCATATAACCAAGACAGATATGT 180
QY	244 TTAGCAAGCACTGCTATGAGACTGAGATTGTGTGGCCAAACCAAACTTGCCATGGC 303
DB	181 CTAAACAGCACTGCTATGAAGACTGAAATTTGTGTGGCCAAAGCCAACTTGCCATGGC 240
QY	304 ACTTCCAGTATGATTGTGCCCAAGCAACGGAAACTCTCAGCAAGCTATGAAAAGGAAAG 363
DB	241 ACTTCCAGCATGATTGTGCCCAAGCAACGGAAACTCTCAGCAAGCTATGAGAAGAAAG 300
QY	364 GAACCTGTGTGTAATACCTTTAGCGAGTGTGTCAGAGTCAGATCAAGTGGAAATTTGTGAA 423
DB	301 GAGCTGTGTCAAGTATTTAGCAGTGTGTCAGAGTCGATCAAGTGGAAATTTGTAGAA 360
QY	424 CATCTTATATCCCAATGTGTCAATTACCAACATGGCCATATAAACTCTGTATCTTAAACCT 483
DB	361 CACCTTATATCCCAATGTGTCAATTACCAAGCATGGGCATCAACTCTACTTAAACCT 420
QY	484 ATGTTCCAGAGAGATTTCATAAATCTGCTGCCAGCTCGGGATTGGATCATATCGCTGAG 543

DB	421 ATGCTCAGAGGAGATTTCATAACTGCACCTGCCAGCACCGGGTCTTCGACACACATCGCTGAG 480
QY	544 AACATTCTGTCTACCTGGATGCCAATCACTATGTGCTGCTGAACTTGTGTGCAAGAA 603
DB	481 AACATTCTGTCTACCTGGATGCCAATCACTATGTGCTGCTGAACTTGTGTGCAAGAA 540
QY	604 TGGTACCGAGTACCTCTCTGATGTCATGCTGTGTGGAAGAGCTTATCGAGAGATGTGTGAG 663
DB	541 TGGTACCGGTGACCTGCGAGCGCATGCTGTGTGGAAGAGCTTATCGAGAGATGTGTGAG 600
QY	664 ACAGATTCTCTGTGAGAGAGCGCTGGCAGAACCAAGAGGATGGGACAGTATTTATCAAA 723
DB	601 ACGGACTCTCTGTGCGAGCGCTGGCAGAACCAAGAGGCTGGGACAGTACTTATTCAAA 660
QY	724 AACAAACCTCTCACGGGAATGCTCTCCAACTCTTTTATAGAGACTTTTATCCTAAA 783
DB	661 AACAAACCTCTGATGAGAACGCTCTCCAACTCTTTTATAGAGCGCTTTATCCTAAA 720
QY	784 ATTATACAGACATTTGAGACAATAGATCTAATTGGAGATGTGGAGACATAGTTTACAG 843
DB	721 ATCATACAGACATTTGAGACAATAGATCTAATTGGAGATGTGGGACATAGTTTACAG 780
QY	844 AGAATTCACTGCCGAAGTGAACAAGCAAGAGGATTTACTGTTTACAGTATGATGATCAG 903
DB	781 AGAATCCACTGCCGAGTGAACAAGTAAAGGGGTTTACTGTTTACAGTATGACGACACAG 840
QY	904 AAAATAGTAGCGGCTTCGAGACACACAATCAAGATCTGGGATAAAACACATTTGGA 963
DB	841 AAGATAGTACGCGGCTTCGAGACACACAATCAAGATCTGGGATAAAAGACATCTGGA 900
QY	964 TGAAGGGAATTTCTCACAGGCCATACAGGTTCACTCTCTGCTCCAGTATGATGAGAGA 1023
DB	901 TGAAGCGGATTTCTACGGGCCACACGGGCTCGTCTCTGCTGCTGACATGAGAGG 960
QY	1024 GTGATCATACAGGATCATCGGATTTCAAGCTCAGAGTGTGGAGTGAATAACAGGTGAA 1083
DB	961 GTGATCATACAGGCTCTCAGACTCCACCGTCAGAGTGTGGAGTGAATAACAGGTGAG 1020
QY	1084 ATGCTAAACAGCTTGATTCAACATTTGAAAGCAGTTTCTGCACTTGCGTTTCAATATGCG 1143
DB	1021 ATGCTAAACAGATTGATTCAACATTTGAAAGCAGTTTCTGCACTTGCGTTTCAATATGCG 1080
QY	1144 ATGATGTGACCTGCTCCAAAGATCGTTTCATTTGCTGTATGGGATATGGCTCCCAACT 1203
DB	1081 ATGATGTGACCTGCTCCAAAGACCGTTCCATTCCTGCTGTGGGATATGGCTCCCAACT 1140
QY	1204 GACATTACCTCCGAGGGTGTGCTGGACACCGAGCTGCTCTCAATGTTGTAGACTTT 1263
DB	1141 GACATACCTCCAGAGGGTGTGCTGGGACACCGAGCTGCGGTCAATGTTGTAGACTTT 1200
QY	1264 GATGACAGTACATTTGTTCTGCAATCTGGGGATAGAACTATATAAGGTATGGAAACAAAT 1323
DB	1201 GATGACAGTACATTTGTTCTGCTCTGGAGATAGAAACCATATAAGGTATGGAAACAAAT 1260
QY	1324 ACTTGTGAATTTGAAAGACTTAAATGGACACAAACAGGCAATTCCTGTTTGGAGTAC 1383
DB	1261 ACCTGTGAATTTGTAAGACCTTAAATGGGACACAGGTGGCATCCCTGTTTGGAGTAC 1320
QY	1384 AGGACAGGCTGTGTGAGTGTGCTCATGTGACAAACACTATCAGATATTTGGGACATAGAA 1443
DB	1321 AGAGACAGGCTGTGTGAGTGTGCTCATGTGACAAACACTATCAGATATTTGGGACATAGAG 1380
QY	1444 TGTGTCATGTTTACGAGTGTAGAGGCCATGAGGAATTTGCTGGGTTGATATTCGATTT 1503
DB	1381 TGTGAGCATGCTTCCGAGTGTGGAGGCCATGAGAGTTGGTACGCTGCATTCGATTT 1440
QY	1504 GATAACAGAGATAGTACGTGGGGCTATGATGGAAAAATTAAGGTGTGGATCTTTGTG 1563
DB	1441 GATAACAAAGGATAGTACGTGGGAGCTATGATGGGAAAAATTAAGGTGTGGATCTTTATG 1500
QY	1564 GCTGCTTTGGACCCCGCTCTCTGACGAGGACACTCTGCTACGAGCCCTTGTGGAGCAT 1623
DB	1501 GCTGCTTTGGACCCCGCTCTCTGACGAGGAGACTCTCTGCTGCGGACACTTGTGGAGCAT 1560

Qy	1624	TCGGAGAGATTTTTCGACTACAGTTTGATGTAATTCAGATGTCAGTAGTTCACATGAT	1683
Db	1561	TCTGGAAGAGTTTTCGCTCTCCAGTTTGATGTAATTCAGATGTCAGTAGTTCACATGAT	1620
Qy	1684	GACACATCCTCATCTGGGACTTCCCTAAATGATCCAGCTGCCAGCTGACACCCCGCT	1743
Db	1621	GACACATCCTCATCTGGGACTTCCCTAAATGATCCAGCTGCCAGCTGACACCCCGCT	1680
Qy	1744	TCCCTTCTCGAATACACCTACATCTCCAGATAAATAACCATACACTGACCTCATACT	1803
Db	1681	TCCCTTCTCGAATACACCTACATCTCCAGATAAATAACCATACACTGACCTCATACT	1740
Qy	1804	TGCCAGGACCCATTAAAGTGGGTATTTAACTGTCGCAATACACAGATGACGAC	1863
Db	1741	TGCCAGGACCTGTTAATGTTGCGATTTTAACTGTCGCAATACACAGATGACGAC	1797
Qy	1864	AACAGTAACAACTACTGCCAGTTTTCCTG--GACTAGCCGAGGAGCAGGCTTT	1921
Db	1798	-----AACAACTCTACTCCCGATTCCCGAGCGATGAGCGAGGAGCGGCTTT	1851
Qy	1922	GAGACTCCTTTCGGACACAGTTGCTCGAGTCGCGCCGACAGCGCTACTCAGACAA	1981
Db	1852	GAGACTCCTTTCGGACACAGTCGGTCAGCAG--CGACGAGGAGCGCTGCTCGGC--A	1907
Qy	1982	CTGACTGCTTCAGTGTCTATCAGAAAGATGCTTCTATCAATTTGTAATGATGGAACT	2041
Db	1908	CCGCTGCTTCAGTGTCTATCAGAAAGATGCTT-TATCTTGTGTGAATGATGGAACT	1966
Qy	2042	TTTAAACCTCCCTCCTCC--TCTTTTCACTCTGACCTAGTTTTCCTATT--GG	2097
Db	1967	TTCAAGCTCCCTCCTCCCTTCCCTTCCCTTCCCTGACCTGTTCTTCCCATTTGG	2026
Qy	2098	TTCAGACAAAGTGACTTATAAATATTTAGTGTGTTTTCGACAAAAA	2151
Db	2027	TTCAGACAAAGTGACTTATAAATATTTAGTGTGTTTACAAAAA	2080
RESULT 10			
AF10396			
LOCUS	AF10396	1979 bp mRNA linear ROD 28-JUN-1999	
DEFINITION	Mus musculus beta-transducin repeat-containing protein mRNA, complete cds.		
ACCESSION	AF10396		
VERSION	AF10396.1	GI:5230821	
KEYWORDS	Mus musculus (house mouse)		
SOURCE	Mus musculus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	Winston, J., Ellledge, S.J. and Harper, J.W.		
AUTHORS	1 (bases 1 to 1979)		
TITLE	Direct Submission		
JOURNAL	Submitted (01-DEC-1998) Biochemistry, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
FEATURES	Location/Qualifiers		
source	1..1979		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/db_xref="taxon:10090"		
CDS	1..1710		
	/note="beta TRCP"		
	/codon_start=1		
	/product="beta-transducin repeat-containing protein"		
	/protein_id="AAD41025.1"		
	/db_xref="GI:5230821"		
	/translation="MDPAEVLQEKALFPMNSREDQNGEPKRIPEKNSLRQV NSCARLNGQETVLTSTAMKTECNVAKKLANCTSSMIVPKQKLSAYEKELCV KYEQMSQDVEFVHLSQMHYOHGHSYLRPMLQDFITAPARGLDHITENI LSYLDKSLCAEVLKQWVRYTSDGLWKKLIERWRTSLRWGLPABRWGGYLFK NKPPDENANSPYRALYPKIIQDIETESNWRGHSIQRIHCRSETSKGVCLQYDQKIVSGLRNTIKIWDKSTLECKRLITGHTGSLVCLQYDERVITGSDSTVRWDV NAGEMLDLTHHCEAVLHLRFNNGMVTCSKORSIAVWDMASPTDITLRLRVLUGHRA		

BASE COUNT	535 a	480 c	516 g	448 t
ORIGIN				
Query Match	72.94;	Score 1567.2;	DB 10;	Length 1979;
Best Local Similarity	88.54;	Pred. No. 0;		
Matches 1761;	Conservative 0;	Mismatches 213;	Indels 16;	Gaps 5;
Qy	70	ATGGACCCGCGGAGCGGTCTGCAAGAGAAAGGCACTCAAGTTTATGATTCCTCAGAG	129	
Db	1	ATGGACCCGCGGAGAGCGGTCTGCAAGAGAAAGCGTTAAGTTTATGATTCCTCAGAG	60	
Qy	130	AGAGAAGCTCTTAATAATGGCGAACCCCTAGGAAGATAATACAGAGAGAAATTCATT	189	
Db	61	AGAGAAGCTCTTAATAATGGCGAACCCCTAGGAAGATAATACAGAGAGAAATTCATT	120	
Qy	190	AGACAGACATAACACAGCTGCGCAGACTCTGTTAAACCAAGAAACAGATGTTTAGCA	249	
Db	121	AGACAGACTTAACAAGCTGTCAGAGCTTTGATTAACCAAGAGAGAGATGTTTAGCA	180	
Qy	250	AGCACTGCTATGAAGACTGAGAAATTTGTGGCCAAACAAACTTCCCAATGSCACTTCC	309	
Db	181	AGCACTGCTATGAAGACTGAAATTTGTGGCCAAACAAACTTCCCAATGSCACTTCC	240	
Qy	310	AGTATGATTGTGCCCAAGCAAGGAACTCTCAGCAGCTATGAAAGGAAAGGAACTG	369	
Db	241	AGCATGATTGTGCCCAAGCAGCGAACTCTCAGCAGCTATGAGAGGAAAGGAGCTG	300	
Qy	370	TGTGTCAAAATCTTTGAGCAGTGTGTAGAGTCAGATCAAGTGAATTTGTGGAACTCTT	429	
Db	301	TGTGTCAAAATCTTTGAGCAGTGTGTGGAGTCTGATCAAGTGAATTTGTAGAACCTT	360	
Qy	430	ATATCCCAAAATGTGTCATCCAACTGGCAGACATAAATCTGATCTTAAACCTATGTTG	489	
Db	361	ATATCCCAAAATGTGTCATCCAGCATGGGCACTCACTCTACCTAAACCTATGCTG	420	
Qy	490	CAGAGAGATTTCAATACCTGCTCTCCAGCTCGGGGATTTGGATCATATGCTGAGAACATT	549	
Db	421	CAGAGAGATTTCAATACCTGCTCTCCAGCTCGGGGATTTGGATCATATGCTGAGAACATT	480	
Qy	550	CTGTCTACCTGAGTGGCCAAATCACTATGCTGCTGAACTTGTGTGCGAGGAACTGCTAC	609	
Db	481	CTGTCTACCTGAGCGCCAGTCACTGTGTGCTGCTGAGCTCTGTGTGCGAGGAACTGCTAC	540	
Qy	610	CGAGTGACCTCTGATGCGATGCTGTGGAAGAACTTATCGAGAGATTTGTCAGGACAGAT	669	
Db	541	CGCGTGAGCTCCGACGGCATGCTGTGGAAGAACTTATCGAGAGATTTGTCAGGACGGAC	600	
Qy	670	TCTCTGTGAGAGCGCTGGCAGACGAGAGGATGGGACAGTATTTTATTTCAAAACAAA	729	
Db	601	TCTCTGTGCGAGCGCTGGCAGACGCGCAGAGCGCTGGGACAGTATTTTATTTCAAAACAAA	660	
Qy	730	CCTCTGTAGCGGAATGCTCTCCAACTCTTTTATAGAGCACTTTATCTTCAAAATATA	789	
Db	661	CCTCTGTAGTGAAGCGCTCTCTCCAACTCTTTTATAGAGCGCTTTATCTTCAAAATATA	720	
Qy	790	CAAGACATTTAGACAAATAGAGATCTAATTTGGAGATGCGGAAGACATAGTTTACAGAGATT	849	
Db	721	CAAGACATTTAGACAAATAGAGATCTAATTTGGAGATGCGGAAGACATAGTTTACAGAGATT	780	
Qy	850	CACCTCCGAGTGAACCAAGCAAGAGGAGTTTACTGTTTACAGTATGATGATCAGAAATA	909	
Db	781	CACCTCCGAGTGAACCAAGTAAAGGGGTTTACTGTTTACAGTATGATGATCAGAAATA	840	
Qy	910	GTAACGGCTTTCAGACACACAAATCAAGATCTGGGATTAACACATTTGATTCAGAG	969	
Db	841	GTACGGCTTTCAGACACACACCAATCAAGATCTGGGATTAACACATTTGATTCAGAG	900	
Qy	970	CGAATTTTCAGAGGCGCATACAGGTTTCACTCTGCTCTCCAGTATGATGAGAGATGATC	1029	

Db 901 CGGATTCCTCAGGGGCCACACGGGCTCGTCTGTCTGCGAGTACGATGAGGGGTGATC 960
 QY 1030 ATACAGGATCATCGGATCCAGGTCAGAGTGGGATGTAATATACAGTCAATATGCTA 1089
 Db 961 ATACAGGCTCCTCAGACTCACCGTCAGAGTGGGATGTAATATCAGGTGAGATGCTA 1020
 QY 1090 AACACGTTGATTCACCATTTGAAGCAGTCTCTCAGTTCGTTTCAATATGCGATG 1149
 Db 1021 GACACATGACTCACCACCTGGAAGCGGTTCTGCACCTCGCTTCAATATGCGATG 1080
 QY 1150 GTGACCTCTCCAAAGATCGTTCATTCGTATGGATATGCGCTCCCACTGACATT 1209
 Db 1081 GTGACCTGTCTCCAAAGACCGTTCATCGCTGTGGGATATGCTTCCCACTGACATC 1140
 QY 1210 ACCCTCGAGGGTGTCTGCGGACACCGAGCTGCTGTCATATGTTAGACTTTGATGAC 1269
 Db 1141 ACCCTCAGAGGGTGTCTGCGGACACCGAGCTGCGTCAATGTTGACTTTGATGAC 1200
 QY 1270 AAGTACATTTGTTCTGCTATCTGGGATAGAACTATTAAGGTATGGAACCAAGTCTGT 1329
 Db 1201 AAGTACATCGTTTCTGCTCTGGAGATAGAACCATTAAGGTGTGGAACCAAGTCTGT 1260
 QY 1330 GAATTTCTAAGGACCTTAATATGACACACAAACGAGCAATGCTGTTGAGTACAGGAC 1389
 Db 1261 GAATTCGTAAGACCTTAATATGAGCACAAGCGTGGCATCGCTGTTGAGTACAGAC 1320
 QY 1390 AGGCTGTAGTGTGCTCATCTGACACCACTATCAGATTATGGACATAGAAATGTGT 1449
 Db 1321 AGGCTGTGTGTGAGCGCTCTCTGACACACCATCAGGCTGTGGACATAGAGTGTGA 1380
 QY 1450 GCATGTTTACAGTGTGTAGAGCCATGAGGAATGTGTGTTGATTCGATTGATATAC 1509
 Db 1381 GCATGCTGTGAGTGTGTGAGGCGCATGAGAGTGTGTACGCTGCATTCGATTGATAC 1440
 QY 1510 AAGAGGATGTAGTGGGCTATGATGGAAAAATTAAGTGTGGATCTTGTGGCTGT 1569
 Db 1441 AAGAGGATGTAGGCGACGCTATGATGGAAAAATTAAGTGTGGATCTTATGGCTGT 1500
 QY 1570 TTGACCCCGTGTCTGAGGACGACTCTGTCTAGGACCGCTTGTGAGGATTCGGA 1629
 Db 1501 TTGACCCCGTGTCTGAGGAGGACTCTGTCTGCGGACACTTGTGAGGATTCGGA 1560
 QY 1630 AGAGTGTTCGACTACAGTTTGTGATTAATCCAGATTGTCAAGTGTTCATCATGACACA 1689
 Db 1561 AGAGTGTTCGCTCCAGTTTGTGATTAATCCAGATTGTCAAGTGTTCATCATGACACA 1620
 QY 1590 ATCTCATCTGGGACTTCTTAATGATCCAGTGTCCCAAGCTGAACCCCGCTTCCCT 1749
 Db 1621 ATCTCATCTGGGACTTCTTGAATGATCCAGTGTCTCAGCTGAACCCCGCTTCCCT 1680
 QY 1750 TCTCGAATACACCTACATCTCCAGATTAATTAACCATACACTGACCTCATCTTCCCA 1809
 Db 1681 TCTCGAATACACCTACATCTCCAGATTAATTAACCATACACTGAGCTCATTAATGCCA 1740
 QY 1810 GGACCCATTAAGTGTGGTATTAAGTATTCGCAATTCGCAATACAGGATGAGCAACAGT 1869
 Db 1741 GGACTCGTTAATGTGAGTATTAACAGACCTGCGCAAGACAGGATGAAC----- 1791
 QY 1870 AACATCAACTACTGCCCCAGTTTCCCTG--GACTAGCCGAGGACGAGGCTTGAAGT 1927
 Db 1792 AACATCAACTCTTACCCGATTCCTCGGACGATGAGCGAGGAGGAGGCTTGAAGT 1851
 QY 1928 CCTGTTGGGACACAGTTGTTGCTGAGTTCGCGCCAGGACGCTCTACTCAGCACAACTGAT 1987
 Db 1852 CCTGTTGGGACACAGTCCGTCAGCAG--CCGACCCAGGACGCTCTGCTCCG--ACCGCT 1907
 QY 1988 GCTTCAGTGTCTATCAGAGATGCTTCTTATCAATGTGAATGATGGAATTTTAA 2047
 Db 1908 GCCTCAGTGTCTATCAGAAGATGCTT--TATCTTGTGTGAATGATGGAATTTTAA 1966
 QY 2048 CCTCCCTCC 2057
 Db 1967 CCTCCCTCC 1976

RESULT 11
 AF099932 1712 bp mRNA linear ROD 21-JAN-1999
 LOCUS Mus musculus beta-TrCP protein E3RS-IkappaB mRNA, complete cds.
 DEFINITION AF099932
 ACCESSION AF099932
 VERSION AF099932.1 GI:4008019
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1712)
 Yaron, A., Hatzubai, A., Davis, M., Lavon, I., Amit, S., Manning, A. M.,
 Andersen, J. S., Mann, M., Mercurio, F. and Ben-Neriah, Y.
 Identification of the receptor component of the
 IkappaBalpha-ubiquitin ligase
 Nature 396 (6711), 590-594 (1998)
 99075339
 9859996
 2 (bases 1 to 1712)
 Yaron, A., Hatzubai, A., Mercurio, F., Manning, A. M., Andersen, J. S.,
 Mann, M. and Ben-Neriah, Y.
 Direct Submission
 Submitted (19-Oct-1998) Immunology, Hebrew University of Jerusalem,
 Bin Karem, Jerusalem 91120, Israel
 Location/Qualifiers
 1..1712
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 1..1710
 /note="beta-transducin repeat-containing protein; receptor
 component of IkBa-ubiquitin ligase"
 /codon_start=1
 /product="beta-TrCP protein E3RS-IkappaB"
 /protein_id="A408020.1"
 /db_xref="GI:4008020"
 /translation="MDPAVLQEKALFKMNSRERDCNNGEPKRIPEKNSLRQTY
 NSCARLINCNETVCLTSTAMKTENCVCVAKLANGTSSMIVPKQKLSASYPKEKELCV
 KYFEQWSESDQVEFVEHLTSOMCHYQHGHINSYKPMQLQDFITALPARGHDAENI
 LSYLDAKSLCAELCKEYRVTSQMLWKLIERMVLTQSLRWGLARRGWGQVLPK
 DKPPDENAPNSFYALYPKIIQDIETESNWRCPGHSLOHCESETSKGVCIQVD
 DOKIVSGLENDRIKINDKSTLECKAILGHTGCVLCLOYDSERVITGSSDSFVMDV
 NAGEMLNTLIHCEAVLHFRFNNMMVTCSDRSIAYMDMASPTDITLRLVVGHRAA
 VNVVDDKXIIVSASGDRITIKVWNTSTCEFVRTLNHGRKGIACLOYRDLRVVSSGN
 TIRLMDIEGACILVLEGHEELVRCIRFDNKRIVSGAYDKIKVMDLMAALDPRAPAG
 TILCLRTLVEHSGRVFRLQDFEQIVSSSHDDITILINDFLNDFAAHEPPRPSRTTY
 ISR"
 BASE COUNT 469 a 399 c 453 g 391 t
 ORIGIN
 Query Match 67.5%; Score 1451.2; DB 10; Length 1712;
 Best Local Similarity 90.5%; Pred. No. 0;
 Matches 1549; Conservative 0; Mismatches 163; Indels 0; Gaps 0;
 QY 70 ATGAGCCCGCGAGCGGTGCTGCAAGAGAGGACACTCAAGTTTATGAATTCCTCAGAG 129
 Db 1 ATGAGCCCGCGAGCGGTGCTGCGAGGAGAACGCTTAAGTTTATGAATTCCTCAGAG 60
 QY 130 AGAAGAAGCTGTAATAATGCGAACCCCTCTAGGAAGATAATACCGAGAAGAATTCACCT 189
 Db 61 AGAAGAAGCTGTAATAATGCGAACCCCTCTAGGAAGATAATACCGAGAAGAATTCACCT 120
 QY 190 AGACAGACATACAGAGCTGTGCCAGACTCTCTCTTAACCCAGAAACAGATGTTAGCA 249
 Db 121 AGACAGACTTACACAGCTGTGCCAGCTTTCATATAAACCAAGACAGATGTGCTAACA 180
 QY 250 AGCAGCTGCTATGAAGACTGAGAAATTTGTGTGGCCAAAAACAAACTTGCCTTCC 309
 Db 181 AGCAGCTGCTATGAGACTGAGAAATTTGTGTGGCCAAAGCCAACTTGCCTTCC 240

NAGEMLNTLHCEAVLHFRNNGMMVTCSDRSIAVMDASPTDITLRLVLGHRAA
VNVVDFDDKIVISAGSDRTIKVNTSTCEVFRKTLNKHGKIACIQYDRDLVWSGSDN
TIRLDIEGACLRVLGHEHELRVIRCFDNDKRIVSGAYDGKI KWDJLMAALDRPAPAG
TLCLRLTVHESRFRVLQDFEPIQIVSSSHDDTILLNDPLNPAHAEPSPSPRTYTY
15R"

BASE COUNT 469 a 399 c 453 g 389 t

Query Match 67.1%; Score 1442.8; DB 10; Length 1710;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 1543; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY	70	ATGACCCGCGCAGCGGTGTGCAAGAGAGAGCACTCAAGTTTATGAATTCCTCAGAG	129
DB	1	ATGACCCGCGCAGCGGTGTGTCAGGAGAGAGCGCTTAAGTTTATGAATTCCTCAGAG	60
QY	130	AGAGAGAGCTGTAATTAATGGCAACCCCTAGAGAGATAATACACAGAGAGAAATTCACCTT	189
DB	61	AGAGAGAGCTGTAATTAATGGCGAACCCTTAGAGAGATAATACACAGAGAGAAATTCACCTT	120
QY	190	AGACAGACATACAAACAGCTGTGCCAGACTCTCTCTTAAACCAAGAAACAGTATGTTAGCA	249
DB	121	AGACAGACTTACAAACAGCTGTGCCAGCTTTCATATAACCAAGAGACAGTATGTTAACA	180
QY	250	AGCAGCTGATGAAGACTCAGAAATTTGTGGCCAAACCAAACTTGGCAATGGCACTCC	309
DB	181	AGCAGCTGATGAAGACTGAAATTTGTGGCCAAAGCCAACTTGGCAATGGCACTTCC	240
QY	310	AGTATGATTGTGCCCAAGCAACGGAACCTCTCAGCAAGCTATGAAAAGGAAGAACTG	369
DB	241	AGCATGATTGTGCCCAAGCAGCGGAACTCTCAGCAAGCTATGAAAGGAAGAAAGAGCTG	300
QY	370	TGTGTCAATATCTTTGAGCAGTGTGTCAGAGTCAAGTCAAGTGGAAATTTGTGGAACATCTT	429
DB	301	TGTGTCAAGTATTTTGGCAGTGTGTCAGAGTCTGATCAAGTGGAAATTTGTGGAACACCTT	360
QY	430	ATATCCCAATGTGTCTATACCAACATGGGCAATAACTCGTATCTTAAACCTATGTTG	489
DB	361	ATATCCCAATGTGTCTATACCAAGATGGGCAATCACTCTCTAAACCTATGTTG	420
QY	490	CAGAGAGATTTCAATACCTGTGTCAGCTCGGAGATTGGATCAATCGCTGAGAACAT	549
DB	421	CAGAGGATTTCAATACCTGTGTCAGCTCGGAGATTGGATCAATCGCTGAGAACAT	480
QY	550	CTGTCACTACCTGATGCGCAATCACTATGCTGTGCTGAACTGTGTGGAAGAAATGGTAC	609
DB	481	CTGTCACTACCTGATGCGCAATCACTATGCTGTGCTGAACTGTGTGGAAGAAATGGTAC	540
QY	610	CGAGTCACTCTGATGCGCAATCACTATGCTGTGCTGAACTGTGTGGAAGAAATGGTAC	669
DB	541	CGGTGACCTGCGGACGCGCATGCTGTGGAAGAACTCATCGAGGAGTGTGAGGACGAC	600
QY	670	TCTCTGTGAGAGGCGCTGCGAGACGAGAGATGGGACAGTATTTATCAAAACAAA	729
DB	601	TCTCTGTGAGAGGCGCTGCGAGAGCGAGAGCTGGGACATATTTATCAAAACAAA	660
QY	730	CCTCTGTGAGGGAATGCTCTCCAACTCTTTTATGAGCACTTTATCTTAAATATATA	789
DB	661	CCTCTGTGAGGGAATGCTCTCCAACTCTTTTATGAGCGCTTTATCTTAAATATATA	720
QY	790	CAGACATTTGAGACATATGATTAATTTGAGATGTGGAGACATAGTTTACAGAGAAAT	849
DB	721	CAGACATTTGAGACATATGATTCATTTGAGATGTGGGCGACATAGTTTACAGAGAAAT	780
QY	850	CAGTCCGGAATGGAACAAAGCAAGAGTTTACTGTTTACAGTATGATGATCAGAAAAATA	909
DB	781	CAGTCCGGAATGGAACAAAGCAAGAGTTTACTGTTTACAGTATGATGATCAGAAAAATA	840
QY	910	GTAAGGCGCTTCGAGACACACAACTCAAGATCTGGATATAAAACACATTTGGAATGCAAG	969
DB	841	GTCAGGCGCTTCGAGACACACCACTCAAGATCTGGATATAAAAGCACACTGGAATGCAAG	900
QY	970	CGAATTCACAGGCCATACAGGTTTCAGTCTCTCTCCAGTATGATGAGAGAGTGATC	1029

DB	901	CGAATTCACGGGCCACACAGGCTCCGCTCTCTGTGTCTGAGTACGGTGTGAGAGGTTGATC	960
QY	1030	ATAACAGAGATCATCGGATTCACGGTTCAGAGTGTGGATGTAAATACAGGTGAATGCTA	1089
DB	961	ATCACAGGCTCTCAGACTCCACCGTTCAGAGTGTGGATGTAAATGAGGTGAGATGCTA	1020
QY	1090	AACAGCTTGATTCAACCATTTGTGAAGCAGTTCTGCACCTTTCGCTTTCAATATGGCATGATG	1149
DB	1021	AACACATTGATTCAACCATTTGTGAAGCAGTTCTGCACCTTTCGCTTTCAATATGGCATGATG	1080
QY	1150	GTGACCTGCTCCAAAGATCGTTCCATTGCTGTATGGGATATGGCTCCCCAACCTGACAT	1209
DB	1081	GTGACCTGCTCCAAAGACCGTTCCATTGCTGTATGGGATATGGCTCCCCAACCTGACAT	1140
QY	1210	ACCCTCGGAGGCTGCTGGTCCGACACCCAGCTGCTGTCAATGTTGTAGACTTTGATGAC	1269
DB	1141	ACCCTCAGAGGCTGCTGGTGGGACACCCAGCTGCTGTCAATGTTGTAGACTTTGATGAC	1200
QY	1270	AAGTACATTTTCTGCACTCTGGGATAGAACTATAAGGTATGGAAACAAAGTACTTGT	1329
DB	1201	AAGTACATCGTTTCTGCTCTGGGATAGAACTATAAGGTATGGAAACAAAGTACTTGT	1260
QY	1330	GAAATTTGAAGACCTTAAATGGACACAAACGAGCATTGCTGTTTGCAGTACAGGAC	1389
DB	1261	GAAATTCGTAAGACCTTAAATGGGCAAAAGCTGCTGCTGTTTGCAGTACAGGAC	1320
QY	1390	AGGCTGTAGTGAAGTGGTCTATCTGACAACTATCAGATTATGGGACATAGAAATGTTGT	1449
DB	1321	AGGCTGTGTTGAGGCTCTCTCTGACAACTATCAGCTGTGGGACATAGAGTGTGGA	1380
QY	1450	GCATGTTTACAGTCTTAGAGGCGCATGAGGAATTTGCTGCTGTATTCGATTTGATAAC	1509
DB	1381	GCATGCTCGAGTGTGAGGCGCATGAGGAGTTTGTACCTGCAATTCGATTTGATAAC	1440
QY	1510	AGAGGATAGTCAAGTGGGCTTATGATGAAATAATTAAGTGTGGGATCTTGTGGCTGCT	1569
DB	1441	AAAAGGATAGTCAAGTGGGCTTATGATGAAATAATTAAGTGTGGGATCTTGTGGCTGCT	1500
QY	1570	TTGGACCCCGTGTCTGTCAGGAGACATCTGTCTACGGACCTTGTGGACATTCGGGA	1629
DB	1501	TTGGACCCCGTGTCTGTCAGGAGACATCTGTCTGTCGGGACATTTGTGGACATTCGGGA	1560
QY	1630	AGAGTTTTTCGACTACAGTTTATGATTAATTCAGATTGTCACTAGTTCACATGATGACACA	1689
DB	1561	AGAGTTTTTCGCTCCAGTTTATGATTAATTCAGATTGTCACTAGTTCACATGATGACACA	1620
QY	1690	ATCCTCATCTGGGATCTTCTAAATGATCCAGTCCCGACCTGAAACCCCGCTTCCCT	1749
DB	1621	ATTCTCATCTGGGATCTTCTGAAATGATCCAGTCTCAGCTGAACCCCGCTTCCCT	1680
QY	1750	TCTCGAATACATACCTACATCTCCAGATAA	1779
DB	1681	TCTCGAATACATACCTACATCTCCAGATAA	1710

RESULT 13			
XELSBTRCP			
LOCUS	1671 bp	mRNA	linear
DEFINITION	African clawed frog beta-TxCP mRNA, complete cds.		
ACCESSION	M98268		
VERSION	M98268.1	GI:295542	
KEYWORDS	beta-transducin repeats.		
SOURCE	Xenopus laevis (African clawed frog)		
ORGANISM	Xenopus laevis		
REFERENCE	1 (bases 1 to 1671)		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;		
	Xenopodinae; Xenopus.		
	Specvak.W., Keiper,B.D., Stratowa,C. and Castanon,M.J.		
	Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in		
	anaphase are rescued by Xenopus cDNAs encoding N-ras or a protein		
	with beta-transducin repeats		

JOURNAL Mol. Cell. Biol. 13 (8), 4953-4966 (1993)
 MEDLINE 9330289
 PUBMED 8391141
 COMMENT Original source text: Xenopus laevis (library: S. cerevisiae expression library of X.laevis oocytes) cDNA to mRNA.
 FEATURES
 source
 1. 1671
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /cell_type="oocyte"
 /tissue_lib="S. cerevisiae expression library of X.laevis oocytes"
 gene
 1. 1671
 /gene="beta-TrCP"
 CDS
 48..1604
 /gene="beta-TrCP"
 /codon_start=1
 /protein_id="AAA02810.1"
 /db_xref="GI:295543"
 /translation="MEGFSCSLQPPPTASERDCNRPPEPRKILITEKNTURQKLANGT
 SSMIVPKORKLSANKEKELVKYFQWSECDQVEFVHLISRMCHYOHGHINTYLA
 PMLORDETALPARGLDHIAENILSLYLDKSLCSALVCKEWRVTSQMLWKLLER
 MYRDSLRWGLAERKMGQYLFKNKPPDGKTPNSFYRALYFKAIODIETIESNRQCG
 RHLQRIHCRSETSGVYCLQYDDQKIVSGLRNTIKIWDKMTLECKRVLNMGHTGVSIL
 CLOYDERVIITSSDSIVRVNVTGEMNLTHHCEAVLHLRFNNMVTCSKDRSI
 AVMDYATDITLRLVGHRAAVNVDFDKYIVASGDRITKVMNTTCFVFETSLN
 GHKRGICLQVDRLVVSGSDNTIRLWDIECGACLRVLEHGEELVRCIRFONKRIVS
 GAYDGKIKVMDLVAALDPRAPAGTLCRLTVLHSGRVFRFLQDFEIVSSSHDDTILI
 WDFLNDPLGA"
 polyA_site
 1671
 /gene="beta-TrCP"
 BASE COUNT 495 a 341 c 424 g 411 t
 ORIGIN
 Query Match 47.3%; Score 1017.8; DB 5; Length 1671;
 Best Local Similarity 81.8%; Pred. No. 2e-277;
 Matches 1175; Conservative 0; Mismatches 262; Indels 0; Gaps 0;
 QY 284 AAACAAACTTCCGCACTCCAGTATGATGTGCCAAGCAAGCAAGCAACTCTGAG 343
 DB 157 AGACAAACTTCCGCACTCCAGTATGATGTGCCAAGCAAGCAAGCAACTCTGAG 216
 QY 344 CAAAGTATGAAAGGAAAGCACTGTGTGTCATTAATCTTTGAGCAGTGTGAGCTCAG 403
 DB 217 CAAATACGAGAGGAAAGCACTGTGTGTCATTAATCTTTGAGCAGTGTGAGCTGCG 276
 QY 404 ATCAAGTGAATTTGGAAACATCTTATATCCAAATGTGTCAATACCAACATGGGCACA 463
 DB 277 ATCAAGTAGAGTTTGTGAACACCTGATATCTCGAATGTGCCACTATCGACATGGACATA 336
 QY 464 TAAACTCGTATCTTAAACCTATGTTGACAGAGAGATTTTCAATCTGTCTGCCAGCTCGGG 523
 DB 337 TAAACACTTACCTTAAGCAATGTTACAAAGGATTTTCAATACCGCACTACCACTCGCG 396
 QY 524 GATTCGATCATCTCTGAGAACATCTGTCTATACCTACCTGGATGCAATCAATCAATGTGCTG 583
 DB 397 GACTCGATCATACAGAAATAATATCTTTTCAATACCTGGATGCAATCAATCAATGTGCTG 456
 QY 584 CTGAACCTGTGTGCAAGGAATGTTACCGAGTGACCTCTGTATGGCATCTGTGGAAGAAGC 643
 DB 457 CAGAACTGGTATTAAGAGTGTGTATCGAGTGACCTCGATGGAAATGCTCTGGAAGAGC 516
 QY 644 TTATCGAGAGATGTCTAGGACAGATTTCTGTGGAGAGCGCTGGCAAGACGAGAGGAT 703
 DB 517 TCATAGAGCGGATGTCTGGGACAGATTTCTTTTGGAGAGGAGCTGGCAGAAAGAGAGGAT 576
 QY 704 GGGGACAGTATTTTCAAAACCAACCTCTCTGAGGGGAAATGTCCTCCCAACTCTTTT 763
 DB 577 GGGGCAATATTTTGTAAACCAACCTCTCAGATGGGAAACGCCCAACCAATCTTCT 636
 QY 764 ATAGAGCACTTTATCTCTAAATATATACAGACATTTGAGACATAGATCTTAATTTGGAGAT 823

Db 637 ACAGAGCGCTTTACCCAAAAATTATTCAGACATAGACAAATCGAGTCCAACTCGCGCT 696
 QY 824 GTGGAAGACATAGTTTACAGAGAAATTCACCTCCGAGTGAACAAAGCAAGAGGATTTACT 883
 Db 697 GTGGGAGACACAGCTTACAAAGAAATTCACCTCCGAGTGAACAAAGCAAGAGGATTTACT 756
 QY 884 GTTTCAGTATGATGATCAGAAATAGTAAAGGGGCTTTCGAGACAAACACATCAAGATCT 943
 Db 757 GTCTGCAGTATGATGATCAGAAATAGTAAAGGGGCTTTCGAGACAAACACATCAAGATCT 816
 QY 944 GGGATAAAAACACATTTGGAATCAAGCGAATTTCTACAGGCCCATACAGGTTTCACTCTCT 1003
 Db 817 GGGATAAAGAAATCTTTGGAGTCAAGCGAGTGTCTGATGGGTCAACATGGGTCTCTCT 876
 QY 1004 GTCTCCAGTATGATGAGAGAGTATCAAAACAGGATTCATCGATTCACCGTCCAGGATCT 1063
 Db 877 GTCTGCAATATGATGAGAGAGTATCAATCTTGGGTCTTTCAGATCTTACCGTCCGAGTGT 936
 QY 1064 GGGATTAATACAGGTGAATGCTAAACACAGTGTATTCACCATTTGTGAAGCAGTTCTGTC 1123
 Db 937 GGCACGTGAACACAGGAGAAATGTTTGAACACAGCTGATTCACCATCTGTGAGCTGTGCTGC 996
 QY 1124 ACTTGGTTTCAATANTGSCNTGATGGTGACCTGTCTCCAAAGATCGTTCGATTCGCTGAT 1183
 Db 997 ACTTGGAGTTTAAATATGTCATGATGGTGACCTGTCTCCAAAGATCGTTCGATTCGATTT 1056
 QY 1184 GGGATATGCTCTCCCACTGACATTAACCTCCGAGGAGTGTCTGTCGACACCGAGCTG 1243
 Db 1057 GGCACATGSCCTCTGCGACCGATATCACATTACGAAGAGTTTCTGGTAGGCCACCGAGCTG 1116
 QY 1244 CTGTCATGTTGTAGACTTTGATGCAAGTACATTTGTTCTGCACTCGGGATAGACTA 1303
 Db 1117 CCGTAAATGTGGTGGACTTTGATGCAAGTATATAGTCTCTGCACTCTGGTATCGAACA 1176
 QY 1304 TAAAGGTATGGAACACAACTACTTGTGAATTTGTAAGGACCTTAAATGAGACACAAACGAG 1363
 Db 1177 TAAAGGTTTGGAAACACCACTATCATGTGAATTTGTGCGGACATTTGAAACGCGCCACAAAGCGT 1236
 QY 1364 GCATTCGCTTGTGAGTACAGGACAGCTGTGATGAGTGGCTCATCTGACAACTA 1423
 Db 1237 GTATTCGATGCTTACAGTATCGAGATCGGCTTGTGAGTGGTCTTCTGACAAACAA 1296
 QY 1424 TCAGATTTAGGACATAGAAATGTGTGATGTTTACGAGTGTGTAAGAGCCATCAGGAAT 1483
 Db 1297 TCAGATTTGGGATATTGAATCGGTGATGTTTCCGGGTTCTGGAAGGACATGAAGAAC 1356
 QY 1484 TGGTGGTGTATTTGATTTGATTAACAGAGATGATCAGTGGGCTTATGATGAAAAA 1543
 Db 1357 TGGTCCGGTGCATCCGCTTTGATTAACAGAGAAATAGTCAGTGGAGCATATGACGGGAAA 1416
 QY 1544 TTAAGTGTGGGATCTTGTGGTGTCTTTGGACCCCGCTGCTTCGAGGACACTCTGTC 1603
 Db 1417 TTAAGTGTGGGACCTTGTGCTGCTTTGGACCCCGCTGCTTCGAGGACCTGTC 1476
 QY 1604 TAGGACCTTGTGGAGCAATCCGGAAGAGTTTTCGACTACAGTTCGATGAAATTCAGA 1663
 Db 1477 TCCGGACTCTTGTGGAGCAATCCGGAAGAGTTTTCGAGTGTGAGTTCGAGTTCGAGA 1536
 QY 1664 TTGTCAGTATGTCATGATGACAAATCTCTCATCTGGGACTTCTTAATGATCCAG 1720
 Db 1537 TAGTAAGCAGCTTCTCAGATGACATATCTCTCATCTGGGATTTTCTCAACGACCCAG 1593

RESULT 14
 AB033279
 LOCUS
 DEFINITION Homo sapiens BTRCP2 mRNA for F-box and WD-repeats protein
 beta-TRCP2 isoform A, complete cds.
 ACCESSION AB033279
 VERSION AB033279.1
 KEYWORDS BTRCP2; F-box and WD-repeats protein beta-TRCP2 isoform A.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens


```

2Y 1738 CCCGTTCCCTTCTCGAACATACACCTACTCTCCAGATTAATACCATACACTGACC 1796
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1587 ACCGGTCTCCCTCCAGAACATACACTTACTCTAGATTAACAGTCTGCACCTTCCACC 1645

RESULT 15
LOCUS AB033280 2252 bp mRNA linear PRI 14-MAR-2000
DEFINITION Homo sapiens BTRCP2 mRNA for F-box and WD-repeats protein
ACCESSION AB033280
VERSION beta-TRCP2 isoform B, complete cds.
KEYWORDS BTRCP2; F-box and WD-repeats protein beta-TRCP2 isoform B.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Koike,J., Sagara,N., Kirikoshi,H., Takagi,A., Miwa,T., Hirai,M. and
Katoh,M.
Molecular cloning and genomic structure of the betaTRCP2 gene on
chromosome 5q35.1
JOURNAL Biochem. Biophys. Res. Commun. 269 (1), 103-109 (2000)
MEDLINE 20160458
PubMed 10594485
REFERENCE 2 (Bases 1 to 2252)
AUTHORS Katoh,M.
DIRECT SUBMISSION
TITLE Submitted (05-OCT-1999) Masaru Katoh, National Cancer Center,
Genetics Division; Tsukiji 5-chome, Chuo-ku, Tokyo 104-0045, Japan
(E-mail:mkato@ncc.go.jp, Tel:81-3-3542-2511(ex.4402),
Fax:81-3-3541-2685)
FEATURES
    Location/Qualifiers
        1..2252
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /chromosome="5"
            /map="5q35.1"
            /tissue_type="fetal lung"
        1..2252
            /gene="BTRCP2"
            /locus_tag="BTRCP2"
            /gene="BTRCP2"
            /codon_start=1
            /product="F-box and WD-repeats protein beta-TRCP2 isoform
            B"
            /protein_id="BAA92330.1"
            /db_xref="GI:7209811"
            /translation="MEPDSVIEDKTIELMNTSYMEDQNEDESPKKNLWQISNGTSSV
            IVSRKPSGEYQKEKDLICIKYFDQWSESDQVEFVHLISRMCHVQHGHINSYLPML
            ORDFTLPEOGLDHAENILSYLDARSLCAELVKQWQVISEGMLWKLEIRWVR
            TDLWKLSEBGGWQVLFKNRPDGPNSFVSLPKLIQDILETIESNWRGRENLO
            RLOCSNSKGVCLQYDDEKILSGLDNSIKIWDKTSLECLVLTGHGVSULCLOYD
            ERVIVTSSDSVTVWVWNTGEVNTLIHNEAVLHLRFSNGLMVTSDRSIAWDM
            ASADITLRLVLVGHRAVNVDPFDKYIVSASGDRTIKWSTSTCFVFTILNGHRG
            KIKWMDLQALDPRAPASTLCRLTVLHVSGRVFLQDFBQFISSSHDUILIWDFLN
            VPPSAQNETRSPSTYTVISR"
        598 a 501 c 558 g 595 t
    ORIGIN
        Query Match 41.1%; Score 883.4; DB 9; Length 2252;
        Best Local Similarity 74.8%; Pred. No. 3.2e-239;
        Matches 1122; Conservative 0; Mismatches 371; Indels 6; Gaps 1;

2Y 298 AATGCACCTCCAGTATGTCGCCAAGCAACGGAACCTCTCAGCAAGCTATGAAG 357
Db 271 AATGAACATCATCTGTGATCTCTCCAGAGAGGCCATCAGAGGAACATCAAAA 330

358 GAAAGGAACGTGTGTCAAACTCTTGGCAGTGGTTCAGAGTCAAGTCAAGTGAATTT 417
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 331 GAAAAAGACTTGTGTAATTAATTTTGGACAGTGGTCTGTAATCAGATCAAGTGAATTT 390
QY 418 GTGGAACATCTTATATPCCAAATGTGTCAATACCAACATGGGCACATAAACTCGTATCTT 477
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 391 GTGGAACATCTTATTTTCAGAAATGTGTCAATATCAGCATGGACATATTAATCTTACCTG 450
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 478 AAACCTATGTTGCAGAGAGATTTCACTACTGCTGCCAGCTCGGGGATTCGATCATATC 537
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 451 AAGCCCATGTTGCAGGGGACCTTTATACCGCTTTACAGAGCAAGGCTTAGATCATATA 510
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 538 GCTGAGAACATCTGTCTATPACCTGGATGCCAAATCACTATGTGTCTGTGTAACCTTGTGTC 597
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 511 GCAGAAAAACATCTTTTCGTACCTGGATGCCAGGTCTCTGTGTGCAGCAGAGCTGGTATGT 570
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 598 AAGGAATGTTACCGAGTACCTCTCATGSCATGCTGTGAGAGAGCTTTATCGAGAGATG 657
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 571 AAGAATGGCAGGGAGTGATCTCAGAAGGATGCTTTGGAAGAGCTGATTTGAACGAATG 630
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 658 GTCAGACAGATCTCTGTGAGAGGCTTGGCAGAACCAAGAGGATGGGACAGTATTTA 717
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 631 GTACGCACTGATCCCTATGGAAGAGCACTTTTCAGAAAGAGAGGGGTGGATCAGTACCTG 690
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 718 TTCAAAAAACAAACCTCTCGACGGGAATGCTCTCCAACTCTTTTATAGAGCACTTTAT 777
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 691 TTTAAAAACAGCCACAGATGGC-----CCTCCAAATTCATTTATAGCTCATTTATAC 744
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 778 CTTAAATTTATACAGACATTTGAGACAATAGATATCTAATTGGAGATGTGGAGACATAGT 837
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 745 CCAAGATTTATCAGGATATAGAGACTATAGAACTTAATCTGGCGGTGTGGACGACACAAC 804
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 838 TTACAGAGAAATTCACCTGCCGAAGTGAACAACAGCAAGAGAGTTACTGTGTTTACAGTATGAT 897
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 805 TTTCAGAGGATTCAGTGGCGCTCTGAAAATAGTAAGGTGTCTACTGTTTACAGTACGAT 864
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 898 GATCAGAAAATAGTAGCGGCTTCGAGACAACACAATCAAGATCTGGATTAATAACACA 957
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 865 GATCAAAAAATTTATCAGTGGCTTAGAGATAATTTCTATTAAGATATGGATTAATAACACG 924
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 958 TTGGAATGCAAGGGAATTTCTCAGAGCCATACAGGTTCAAGTCTCTGCTCCAGTATGAT 1017
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 925 CTGGAATGTTTGAAGTGTAAACAGGACACACAGGCTCTGTCTCTGTCTGCAGTATGAT 984
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1018 GAGAGAGTGCATCAACAGGATCATCGGATTCACCGTTCACCGTTCAGAGTGTGGATGTAATACA 1077
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 985 GAGCGTGTCAATTGTAATCGCTTTTCAGATTCTACGGTGAGAGTGTGGATGTGAACACG 1044
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1078 GGTGAATGCTAAACAGCTGTGATTCACCATTTGTGAGCAGGTTCTGCATTCGGTTCGAT 1137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1045 GGTGAAGTTCTTAACACATTTGATCCACCAATAGGCTGTATTGCACTTACGCTTCAGC 1104
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1138 AATGGCATGATGTCACCTGCTCCAAAGATCGTTCCATTGCTGTATGGGATATGGCTCC 1197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1105 AATGGACTGATGTGACCTGTTTCCAAAGACCGCTCCATTGCTGTGTGGACATGGCTTCT 1164
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1198 CCAACTGACATTAACCTCCGGAGGGTGTGGTTCGACACCGAGCTGTCTCAATGTTGTA 1257
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1165 GCAGCTGACATCACTTTACGCCGTCTCTGTGGTGGCCCGGGTGCCTCAATGTAGTA 1294
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1258 GACTTTGATGACAAGTACATTTGTTCTGCATCTGGGGATAGAACTATATAAGGTATGGAAC 1317
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1225 GACTTTGACGACAAGTACATCGTGTCTGCTCTGTGTCAGACCATCAAAAGCTGGAGC 1284
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1318 ACAAGTACTTGTGAATTTGTAAAGGACCTTTAAATGGACACAACGAGGATTCCTGTTTG 1377
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1285 ACAGGACCTGTGAATTTGTTTCTGTTACTCTCAATGGGCACAAGCGGGGCAATTCCTGTC 1344
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1378 CAGTACAGGACAGGCTGGTGTAGTGGCTCATCTGACAACACATCATCAGATTTATGGAC 1437
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1345 CAGTACAGGATCGCTGGTGTGTAGTGGATCATCAGATATATACATTTAGGCTCTGGAT 1404
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1438 ATAGAATGTTGTCATGTTTACAGTGTTAGAGGCCATGAGGAATTCGTCGTTGAT 1497
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1405 ATTGAATGTGGTCTGTTTAAAGATCTCTAGAGGACATGAAGAAATTCGTCGATGCATC 1464
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

y 1498 CGATTGATACAAAGAGGATAGTCAGTGGGGCCCTATGATGGAAAAATTAAGTGGGAT 1557
b      |||||
b 1465 CGGTTTGATTAACAAGAGGATTTGTCAAGTGGGGCCCTATGATGGAAAAATTAAGTGGGAC 1524
y 1558 CTTGTGGCTGCTTTTGGACCCCGTCTCTGCGAGGGACACTCTGTCTACGGACCCCTTGTG 1617
b      |||||
b 1525 TTGCAAGCTGCTCTTGACCCCTCGAGCCCGCAGCAGCACATTTGTGTTGGCACAATGGTG 1584
y 1618 GAGCATTCCGGAAGAGTTTTTTCGACTACAGTTTTCAGTAATTCAGATTGTCAGTAGTTCA 1677
b      |||||
b 1585 GAACATTCTGACGCTGTTTTCGGCTCCAGTTTGATGAGTTTCAGATCATCAGCAGCTCC 1644
y 1678 CATGATGACRCAATCCTCTCTCTGGGACTTCTTAATGATCCAGCTGCCCAAGCTGACCC 1737
b      |||||
b 1645 CATGATGACACTATTTTGAATTTGGGATTTCTTAATGTTGCTCCAGTGGCCGAGATGAG 1704
y 1738 CCGGTTCCCTTCTCGAACAATACACCTACATCTCCAGATAAATAACCATACACTGACC 1796
b      |||||
b 1705 ACCCGTTCTCCCTCCAGAACATACACTTACATCTCTAGATAACAGTCTGCACCTTTCACC 1763

```

search completed: October 22, 2003, 20:48:52
ob time : 7717 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

M nucleic - nucleic search, using sw model

Run on: October 22, 2003, 18:18:24 ; Search time 4097 Seconds
(without alignments)
12760.295 Million cell updates/sec

Title: US-09-601-168B-1
Perfect score: 2151
Sequence: 1 tgcgttggtcgccgtggc.....gtttgcagaaaaaaa 2151

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estmu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_hic.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_hic.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: em_estom.*
- 16: em_estom.*
- 17: em_gss_hum.*
- 18: em_gss_inv.*
- 19: em_gss_pln.*
- 20: em_gss_vrt.*
- 21: em_gss_fun.*
- 22: em_gss_mam.*
- 23: em_gss_mus.*
- 24: em_gss_pro.*
- 25: em_gss_rod.*
- 26: em_gss_pig.*
- 27: em_gss_vrl.*
- 28: gb_gss1.*
- 29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1591.2	74.0	2897	11 AK032221	AK032221 Mus muscu
2	1557.2	72.4	2970	11 AK083257	AK083257 Mus muscu
3	1263.8	58.8	4502	11 AK052317	AK052317 Mus muscu
4	877.8	40.8	3824	11 AK041532	AK041532 Mus muscu

5	638.6	32.5	988	13	BX398796	BX398796
6	650.6	30.2	709	10	BE782628	BE782628
7	643.4	29.9	776	14	CB244526	CB244526
8	638	29.7	778	13	BQ444188	BQ444188
9	635.2	29.5	652	9	AL044095	AL044095
10	625	29.1	851	13	BQ425358	BQ425358
11	624.8	29.0	973	12	BM465280	BM465280
12	622.2	28.9	887	13	BUS07728	BUS07728
13	616.6	28.7	744	14	CB520936	CB520936
14	613.2	28.5	758	10	BE296484	BE296484
15	608.8	28.3	692	10	BE269297	BE269297
16	603.2	28.0	749	14	CB246650	CB246650
17	589.4	27.4	828	12	B1822845	B1822845
18	586.6	27.3	1147	13	BQ948186	BQ948186
19	580.6	27.0	743	12	BG829037	BG829037
20	579.4	26.9	895	9	AL898993	AL898993
21	574.2	26.7	699	14	CB246901	CB246901
22	566.4	26.3	752	10	BG722472	BG722472
23	561.6	26.1	1136	14	CD500760	CD500760
24	557	25.9	779	14	CD350839	CD350839
25	553.6	25.7	924	14	CA985560	CA985560
26	540	25.1	870	13	BQ463637	BQ463637
27	534.4	24.8	789	10	BG518761	BG518761
28	534.2	24.8	966	14	CA980902	CA980902
29	532.6	24.8	555	14	CA416190	CA416190
30	532.4	24.8	637	12	B1103092	B1103092
31	529.8	24.6	544	9	AL041438	AL041438
32	528.4	24.5	888	12	B1455545	B1455545
33	521.8	24.2	729	14	CB948025	CB948025
34	518	24.1	756	13	BQ460380	BQ460380
35	516.2	24.0	810	12	BM944304	BM944304
36	513.6	23.9	846	9	AL879775	AL879775
37	511.6	23.8	636	13	BU055333	BU055333
38	500.4	23.3	646	10	BG082149	BG082149
39	497	23.1	850	13	BQ771102	BQ771102
40	494.8	23.0	850	14	CA327271	CA327271
41	494.6	23.0	1002	14	CA971789	CA971789
42	492.2	22.9	541	10	BG058605	BG058605
43	488.8	22.7	590	9	AW762522	AW762522
44	488.2	22.7	806	13	BU750387	BU750387
45	488	22.7	488	13	BX099276	BX099276

ALIGNMENTS

RESULT 1	AK032221	2897 bp	linear	HTC 05-DEC-2002
LOCUS	Mus musculus adult male olfactory brain cDNA, RIKEN full-length enriched library, clone:6430504E23 product:beta-transducin repeat containing protein, full insert sequence.			
DEFINITION	AK032221			
ACCESSION	AK032221			
VERSION	AK032221.1	GI:26082733		
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
AUTHORS	Carninci,P. and Hayashizaki,Y.			
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)			
MEDLINE	99279253			
PUBMED	10349636			
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,N., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.			
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)			
MEDLINE	20499374			
PUBMED	11042159			

892 TATGATGATCAGAAATAGTAAGCGGCTTCGAGACAACATCAAGATCTGGGATATA 951
|||
999 TACGACGACCAAGATAGTACGGGCTTCGAGACAACATCAAGATCTGGGATATA 1058
|||
952 AACACATTTGGAATGCAAGCGAATCTCACAGGCCATACAGGTTCAAGTCTCTCTCCAG 1011
|||
1059 AGCACATGGAATGCAAGCGAATCTCACAGGCCATACAGGTTCAAGTCTCTCTCCAG 1118
|||
1012 TATGATGAGAGATGATCATACAGGATCATGATTCACGGCTCAGAGTGGGATGA 1071
|||
1119 TACGATGAGAGGATGATCATACAGGCTCTCACAGTCCACCGTCCAGAGTGGGATGA 1178
|||
1072 AATCAGGTGAATGCTAAACACGTTGATTCACCATTTGGAAGCAGTCTGCATTTGCT 1131
|||
1179 AATGAGGTGAGATGCTAAACACATGATTCACCATTTGGAAGCAGTCTGCATTTGCT 1238
|||
1132 TTCAATATGGCATGATGGTGAAGTCTCCAAAGATGTTCCATTCGTTGTTGGGATAG 1191
|||
1239 TTCAATATGGCATGATGGTGAAGTCTCCAAAGATGTTCCATTCGTTGTTGGGATAG 1298
|||
1192 GCTTCCCAACTGACATTAACCTCCGAGGGTGTGGTGGACACCGAGTGTCTCAAT 1251
|||
1299 GCTTCCCAACTGACATTAACCTCCGAGGGTGTGGTGGACACCGAGTGTCTCAAT 1358
|||
1252 GTTGTAGATTTGATGACAAGTACATGTTTCTGCATCTGGGATGAGAACTATAAGGTA 1311
|||
1359 GTTGTAGATTTGATGACAAGTACATGTTTCTGCATCTGGGATGAGAACTATAAGGTA 1418
|||
1312 TGGACACAAAGTACTTGTGAATTTGTAAGGACCTTAATAGGACACAAACGAGCATTTGC 1371
|||
1419 TGGACACAAAGTACTTGTGAATTTGTAAGGACCTTAATAGGACACAAACGAGCATTTGC 1478
|||
1372 TGTTCAGTACAGGACAGGCTGTAGTGTGCTCATCTGCACACATTCAGATTA 1431
|||
1479 TGTTCAGTACAGGACAGGCTGTAGTGTGCTCATCTGCACACATTCAGGCTG 1538
|||
1432 TGGGACATAGAAATGTGGTGCATGTTTACGAGTGTAGAGGCCATGAGGAATTTGGTGGT 1491
|||
1539 TGGGACATAGAAATGTGGTGCATGTTTACGAGTGTAGAGGCCATGAGGAATTTGGTGGT 1598
|||
1492 TGTATTCGATTTGATAACAGAGGATGATGATGGGCTTATGATGAATAATTAAGTG 1551
|||
1599 TGATTCGATTTGATAACAGAGGATGATGATGGGCTTATGATGAATAATTAAGTG 1658
|||
1552 TGGGATCTTGTGGCTGCTTTGACCCCGTGTCTCTGAGGACACATCTCTCTACGAC 1611
|||
1659 TGGGATCTTGTGGCTGCTTTGACCCCGTGTCTCTGAGGACACATCTCTCTGCGGACA 1718
|||
1612 CTGTGGACATTCGGAGAGTCTTTTGGACATGATTTGATGATTCAGATTTGTCAGT 1671
|||
1719 CTGTGGAGCATCTGGAGAGTCTTTTGGACATGATTTGATGATTCAGATTTGTCAGT 1778
|||
1672 AGTTCAATGATGACACATCTCTGAGATCTTCTTAAATGATTCAGCTGCTCCAGCT 1731
|||
1779 AGTTCAATGATGACACATCTCTGAGATCTTCTTAAATGATTCAGCTGCTCCAGCT 1838
|||
1732 GACCCCGGCTTCTGACATACACCTCATCTCCAGATTAATTAACATACAC 1791
|||
1839 GACCCCGGCTTCTGACATACACCTCATCTCCAGATTAATTAACATACAC 1898
|||
1792 TGACCTCATACTTGGCCAGGACCCATTAAGTGTGGTATTAACGATTTGCAATPAC 1851
|||
1899 TGCCCTCATATTTGCCAGGACTCGTTAATGTTGAGTATTAACAGACCTTCCAGAC 1958
|||
1852 AGGATGACACACAGTACATCAACTACTGCTCCAGTTCCTG--GACTAGCCGAG 1909
|||
1959 AGGATGAC-----AACATCAACTCTTACCCGGATTTCCCGGATGAGCGAG 2009
|||
1910 GAGCAGGGCTTTGAGACTCTGTTGGGACACAGTGTGGTCTGAGTTCGSCCAGGACGGTC 1969
|||
2010 GAGCAGGGCTTTGAGACTCTGTTGGGACACAGTGTGGTCTGAGCAGCCGAGCAGCGCC 2069
|||

1970 TACTCAGCAACACTGACTGCTTCAAGTCTGCTATCAGAGATGCTTCTATCAATGTA 2029
|||
2070 TGCTCGGC---ACCGGCTGCTCAGTCTGCTATCAGAGAGTCTT-TAICTTGTGTA 2125
|||
2030 ATGATTGGAACCTTTTAAACCTCCCTCTCTCC---TCTTTACCTCTGACCTAGTTT 2086
|||
2126 ATGATTGGAACCTTCCAAGCTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTGCTTTC 2185
|||
2087 TTTCCCATTTGGTTCACAGCAAGGAGTACTTATAATATATTTAGTGTGTTGCCAGAA 2143
|||
2186 CTCCCATTTGGGTTTCAGACCAAGATGACTTATAATATATTTAGTGTGTTGCCAGAA 2243
|||

AK083257 2970 bp mRNA linear HTC 05-DEC-2002
Mus musculus adult male hippocampus cDNA, RIKEN full-length
enriched library, clone: C630031K01 product: beta-transducin repeat
containing protein, full insert sequence.

AK083257 1 GI:26101148
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunuma, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, K.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Aizawa, K., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Iizawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gotohori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barthe, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Holman, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Nordone, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaur, P., Sasaki, H.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
and Hayashizaki, Y.


```
1370 CCTGTTTGAGTACAGGACAGCGCTGGTAGTGAGTGGCTCATCTGACCAACACTATCAGAT 1429
1491 CTTGTTTGCAGTACAGACAGCGCTGGTAGTGAGTGGCTCTCTGACCAACACTATCAGC 1550
1430 TATGGACATAGAAATGTGGTGCATGTTTACGAGTGTGTAGAGGCGCATGAGAAATGGTGC 1489
1551 TGTGGACATAGAGTGTGGAGCATGCTCGAGTGTGTGGAGGCGCATGAGAGTGGTATC 1610
1490 GTTGATTCGATTTGATAAAGAGAGATAGTCAGTGGGGGCTATGATGGAAAAATTAAG 1549
1611 GCTGCATTCGATTTGATAAAGAGAGATAGTCAGGCGGCTATGATGGAAAAATTAAG 1670
1550 TGTGGATCTTGTGGCTGTGTGGACCCCGTGTCTCTGAGGACACTCTGTCTACGGA 1609
1671 TGTGGATCTTGTGGCTGTGTGGACCCCGTGTCTCTGAGGAGGACTCTGTCTGCGGA 1730
1610 CCTGTGTGAGCATTTCCGGAAGAGTGTTCGCACTACAGTGTGATGAATCCAGATTGTCA 1669
1731 CACTTGTGGAGCATTTCTGGAAGAGTGTTCGCGCTCCAGTGTGATGAATCCAGATTGTCA 1790
1670 GTAGTTCATGATGACACAAATCTCTCTGGGACTTCTCTAAATGATCCAGCTGCCAAG 1729
1791 GTAGTTCATGATGACACAAATCTCTCTGGGACTTCTCTAAATGATCCAGCTGCCAAG 1850
1730 CTGAACCCCGGCTTCCCTCTCTGACACATACACTACATCTCCAGATAAATAACCATAC 1789
1851 CTGAACCCCGGCTTCCCTCTCTGACACATACACTACATCTCCAGATAAATAACCATAC 1910
1790 ATGACCTCATACTTGTGCGGAGGACCCATTAAGTGGGTATTTAAAGTATCTGCCAATA 1849
1911 ACTGGCTCATTAATTTGCCAGGACTCGTAAATGTGTGAGCATTTTAAACAGACCTGCCAAGA 1970
1850 CAGAGATGAGCAACACAGTAACTAACTACTGCGGACTTCCCTG--GACTAGCGG 1907
1971 CAGAGATGAGCAACACAGTAACTAACTACTGCGGACTTCCCTG--GACTAGCGG 2021
1908 AGAGAGCGGCTTGGAGACTCTCTGTGGGACACAGTGTGTCTGAGTGGCGCCAGGACGG 1967
2022 AGGAGCGGCTTGGAGACTCTCTGTGGGACACAGTGTGTCTGAGTGGCGCCAGGACGG 2080
1968 TCTACTCAGCACAAGTCTGCTCAGTGTGCTGCTATCAGAGAGTCTTTATCAATTGT 2027
2081 CTTGCTCGGC---ACCGGCTGCTCAGTGTGCTGCTATCAGAGAGTCTTTATCAATTGT 2136
2028 GAATGATGGAACTTTTAAACCTCCCTCCCTCTCC---TCCTTTCACCTCTGCACTAGT 2084
2137 GAATGATGGAACTTTTAAACCTCCCTCCCTCTCC---TCCTTTCACCTCTGCACTAGT 2196
2085 TTTTTCCTTCCAT--GGTCCAGACAAAGTGTGCTATTAATATATATTAGTGTTCGCCAGAA 2143
2197 TCCCTCCATGGGTTCAGACAAAGATGACTTATAAATATATTAGTGTTCGCCAGAA 2256

>RESULT 3
>LOCUS AK052317
>DEFINITION Mus musculus 13 days embryo heart cDNA, RIKEN full-length enriched
library, clone:D330027D1 product:beta-transducin repeat containing
protein, full insert sequence.
>ACCESSION AK052317
>VERSION AK052317.1 GI:26095121
>KEYWORDS HTC; CAP trapper.
>SOURCE Mus musculus (house mouse)
>ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
>REFERENCE Carninci, P. and Hayashizaki, Y.
>TITLE High-efficiency full-length cDNA cloning
>JOURNAL Meth. Enzymol. 303, 19-44 (1999)
>MEDLINE 99279253
>PUBMED 10349636
```

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

```
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
>REFERENCE
>AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kusunagi, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, K., Inoue, K., Togawa, Y., Izawa, M., Ohara, S., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
>REFERENCE
>AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, K., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gless, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G.,
Quackenbush, J., Schraml, L. M., Stauble, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bolunga, N.,
Carninci, P., de Bona, M. P., Brownstein, M. J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Nordone, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombauer, P., Nardone, P.,
Ringo, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851
5
>REFERENCE
>AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
>TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
12466851
6 (bases 1 to 4502)
>REFERENCE
>AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koyas, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, E., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahita, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
```


COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>

FEATURES

```

source
    1..4502
        /organism="Mus musculus"
        /mol_type="mRNA"
        /strain="C57BL/6J"
        /db_xref="PANTOM DB:D330027D11"
        /db_xref="taxon:10090"
        /clone="D330027D11"
        /cissue_type="heart"
        /note="beta-transducin repeat containing protein (MGDI|GI:1338871, GB|NM_009771, evidence: BLASTN, 100%, match=1502)"
misc_feature
    1..4502
        /clone_lib="RIKEN full-length enriched mouse cDNA library"
        /dev_stage="13 days embryo"
BASE COUNT      1284 a   998 c   1126 g   1094 t
ORIGIN
Query Match      58.8%; Score 1263.8; DB 11; Length 4502;
Best Local Similarity 89.9%; Pred. No. 0;
Matches 1355; Conservative 0; Mismatches 152; Indels 0; Gaps 0;
2y 112 TTTATCAATTCCTCAGAGAGAGACGTCTAATATGSCGACCCCTAGGAGATAATA 171
Db 280 TTCAGAAATTCCTCAGAGAGAGACGTCTAATATGSCGACCCCTAGGAGATAATA 339
2y 172 CCAGAGAAGAAATTCCTTAGACAGACATACACAGCTGTGCAGACTCTGCTTAAACCAA 231
Db 340 CCAGAGAAGAAATTCCTTAGACAGACATACACAGCTGTGCAGACTCTGCTTAAACCAA 399
2y 232 GAAACAGTAGTTTTCACAGACACTGTATGAAGACTGAGAGTCTGTGGCCAAACAA 291
Db 400 GAGACAGTAGTTTTCACAGACACTGTATGAAGACTGAGAGTCTGTGGCCAAACAA 459
2y 292 CTTGCCAATGGCACTTCCAGTAGTATGTTGCCAAAGCAACGAACTCTCAGCAAGCTAT 351
Db 460 CTTGCCAATGGCACTTCCAGTAGTATGTTGCCAAAGCAACGAACTCTCAGCAAGCTAT 519
2y 352 GAAAGGAAAGAAAGAACTGTGTCAAACTTTGAGCAGTGTGAGAGTCAAGTCAAGTG 411
Db 520 GAGAGGAAAGAAAGAACTGTGTCAAGTATTTTGAAGAGTGTGAGAGTCAAGTCAAGTG 579
2y 412 GAATTTGTGGAACATCTTATATCCCAATGTGTCTATTACCAACATGGGCAATAACTCG 471
Db 580 GAATTTGTGGAACATCTTATATCCCAATGTGTCTATTACCAACATGGGCAATAACTCG 639
2y 472 TATCTTAAACCTATGTCAGAGAGATTTCAATCTGTCTGCCAGTCTGGGGATTGGAT 531
Db 640 TACTTAAACCTATGTCAGAGAGATTTCAATCTGTCTGCCAGTCTGGGGATTGGAT 699
2y 532 CATATCGCTGAGAACATCTGTCTACCTGGATGTCGCAATCACTATGTCGTCTGAACCTT 591
Db 700 CATATCGCTGAGAACATCTGTCTACCTGGATGTCGCAATCACTATGTCGTCTGAACCTT 759
2y 592 GTGTGCAAGGAATGTTACCGAGTGACCTCTGATGGCATGCTGTGAAGAACTTATCGAG 651
Db 760 GTGTGCAAGGAATGTTACCGAGTGACCTCTGATGGCATGCTGTGAAGAACTTATCGAG 819
2y 652 AGAATGTCAGACAGATCTCTGTGGAGAGCCCTGGCAGAACGAGAGGATGGGACAG 711
Db 820 AGAATGTCAGACAGATCTCTGTGGAGAGCCCTGGCAGAACGAGAGGATGGGACAG 879
2y 712 TATTTATTCAGAAACAACTCTCTGACGGGATGCTCTCCCACTCTTTTATAGAGCA 771
Db 880 TACTTATTCAGAAACAACTCTCTGATGAGAACGCTCTCTCCCACTCTTTTATAGAGCG 939

```

```

Qy 772 CTTTATCCTTAAAAATTATAAAGACATTTGAGACAATAGATAATCTAATTGGAGATGTGAAGA 831
Db 940 CTTTATCCTTAAAAATCATAAAGACATTTGAGACAATAGATGTCATTTGGAGATGTGGGCGA 999
Qy 832 CATAGTTTACAGAGAATTCACCTGCCGAAGTGAACAACAAGAGAGGTTTACTGTTTACAG 891
Db 1000 CATAGTTTACAGAGAATTCACCTGCCGAAGTGAACAACAAGAGGTTTACTGTTTACAG 1059
Qy 892 TATGATGATCAGAAAAATAGTAAGCGGCTTCGAGACAACACAATCAAGATCTGGGATAAA 951
Db 1060 TACGACGACCAAGAGATAGTCAGCGGCTTCGAGACAACACCATCAAGATCTGGGATAAA 1119
Qy 952 AACACATTTGGAAATGCAAGCGAATTTCTACAGGCGCATACAGGTTTCTGCTCTCAG 1011
Db 1120 AGCACACTGGAAATGCAAGCGGATTTCTACGGGCGCACAGGGCTCCGTCCTGTCTCTGCG 1179
Qy 1012 TATGATGAGAGATGATCATTAACAGGATCATCGGATTCACGGTCAGAGTGTGGGATGTA 1071
Db 1180 TACGATCAGAGGCGTATCATCAAGGCTCTCAGACTCCACCGTCAGAGTGTGGGATGTA 1239
Qy 1072 AATACAGGTCGAATGCTAAACACGTTGATTCACCATTTGTAAGCAGTTTCTGCACCTGCGT 1131
Db 1240 AATCAGGTCGATGCTAAACACATTTGATTCACCATTTGTAAGCAGTTTCTGCACCTGCGC 1299
Qy 1132 TTCAATATGGCATGATGTGACCTGCTCCAAAGATCGTTCCATTCGTTGTATGGGATATG 1191
Db 1300 TTCAATATGGCATGATGTGACCTGCTCCAAAGACCGTTCCATTCGTTGTATGGGATATG 1359
Qy 1192 GCCTCCCACTACATTTACCTCCGAGGCGTGTGTCGACACCGAGCTGCTGTCAT 1251
Db 1360 GCTTCCCACTACATTTACCTCCGAGGCGTGTGTCGACACCGAGCTGCTGTCAT 1419
Qy 1252 GTTGATAGACTTTGATGACAAAGTACATTTCTGCATCTGGGGATAGAACTATAAAGTGA 1311
Db 1420 GTTGATAGACTTTGATGACAAAGTACATTTCTGCCTCTGGAGATAGAACTATAAAGTGA 1479
Qy 1312 TGGAAACACAGTACTTGTGAATTTGTAAGGACCTTAATATGGACACACAAAGAGGATTCGCC 1371
Db 1480 TGGAAACACAGTACTTGTGAATTTGTAAGGACCTTAATATGGACACACAAAGAGGATTCGCC 1539
Qy 1372 TGTGTCAGTACAGGACAGGCTGTGATGATGAGTGTGCTCATCTGACAAACATATCAGATTA 1431
Db 1540 TGTGTCAGTACAGGACAGGCTGTGATGATGAGTGTGCTCATCTGACAAACATATCAGATTA 1599
Qy 1432 TGGACATACAGATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1491
Db 1600 TGGACATACAGATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1659
Qy 1492 TGTATTCGATTTGATTAACAGAGGATGATGATGATGATGATGATGATGATGATGATGATG 1551
Db 1660 TGTATTCGATTTGATTAACAGAGGATGATGATGATGATGATGATGATGATGATGATGATG 1719
Qy 1552 TGGGATCTTGTGGCTGTGTTGGACCCCGCTGCTCTGACAGGACACTCTGCTACGGACC 1611
Db 1720 TGGGATCTTGTGGCTGTGTTGGACCCCGCTGCTCTGACAGGACACTCTCTGCTACGGACC 1779
Qy 1612 CTTGTGG 1618
Db 1780 CTTGTGG 1786

```

RESULT 4

AK041532

LOCUS

DEFINITION

AK041532 3824 bp mRNA linear HTC 05-DEC-2002
Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched
library, clone:AG30019L11 product:F-BOX/WD-REPEAT PROTEIN 1B (F-BOX
AND WD-REPEATS PROTEIN BETA-TRCP2) homolog (Homo sapiens), full
insert sequence.

ACCESSION

AK041532

VERSION

AK041532.1 GI:26334548

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)


```
256 GCGTGTGATTCGATTTGATTAACAGAGGATAGTCAGTGGGCGCTATGATGGAATAATTAA 315
1548 AGTGTGGGATCTTTGTGGCTGCTTTGGACCCCGTCTCTCCAGGGACACTCTGTCTACG 1607
|||||
316 AGTGTGGGATCTTTGTGGCTGCTTTGGACCCCGTCTCTCCAGGGACACTCTGTCTACG 375
|||||
1608 GACCTTGTGGAGCATTCCGGAAGAGTTTTCGACTACAGTTTGAATGAAATCCAGATTGT 1667
376 GACCTTGTGGAGCATTCCGGAAGAGTTTTCGACTACAGTTTGAATGAAATCCAGATTGT 435
|||||
1668 CAGTAGTTCACATGATGACACAAATCTCATCTGGGACTTCTAAATGATCCAGTGCCCA 1727
436 CAGTAGTTCACATGATGACACAAATCTCATCTGGGACTTCTAAATGATCCAGTGCCCA 495
|||||
1728 AGCTGAACCCCGCTTCCCTTCGGAACATACACTACATCTCCAGATTAATAACCAT 1787
496 AGCTGAACCCCGCTTCCCTTCGGAACATACACTACATCTCCAGATTAATAACCAT 555
|||||
1788 ACACGTACCTCATACTTGGCCAGGACCAATTAAAGTTGCGGTATTTAAAGTATCTGCCAA 1847
556 ACACGTACCTCATACTTGGCCAGGACCAATTAAAGTTGCGGTATTTAAAGTATCTGCCAA 615
|||||
1848 TACCAGGATGACACAAACAGTAACAACTACTGCCAGTTTCCCTGGACTAGCG 1907
616 TACCAGGATGACACAAACAGTAACAACTACTGCCAGTTTCCCTGGACTAGCG 675
|||||
1908 AGGAGCAGGGCT-TTGAGACTCTCTTGGGACACAGTTGGTCTCAGTTCGCGCCAGGACG 1966
676 AGGAGCAGGGCTATTGAGACTCTCTTGGGACACAGTTGGTCTCAGTTCGCGCCAGGACG 735
|||||
1967 GTCTACTAGCACTGACTGCTTCAAGTGTGCTATCAGAAGATGCTTCTATCAATTG 2026
736 GTCTACTAGCACTGACTGCTTCAAGTGTGCTATCAGAAGATGCTTCTATCAATTG 795
|||||
2027 TGAATGATTGGAACCTTTAAACCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2086
796 TGAATGATTGGAACCTTTAAACCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 855
|||||
2087 TTTCCTATGTTTCCAGACAAAGTGACTATAAATATATTAGTGTGTTGCCAGAA 2143
856 TTTCCTATGTTTCCAGACAAAGTGACTATAAATATATTAGTGTGTTGCCAGAA 912
```

```
RESULT 6
3E782628
LOCUS
DEFINITION
601465751F1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3868812 5',
mRNA sequence.
BE782628
1 GI:10203826
ACCESSION
EST.
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9617 row: f column: 13
High quality sequence stop: 655.
Location/Qualifiers
1..709
/organism="Homo sapiens"
```

FEATURES
source

```
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="IMAGE:3868812"  
/tissue_type="retinoblastoma"  
/lab_host="DH10B (phage-resistant)"  
/clone_lib="NIH_MGC_67"  
/notes="Organ: eye; Vector: pCMV-SPORT6; Site: 1: Noif;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.75 kb. Library constructed by Life  
Technologies."  
BASE COUNT 193 a 156 c 175 g 185 t  
ORIGIN
```

```
Query Match 30.2%; Score 650.6; DB 10; Length 709;  
Best Local Similarity 99.2%; Pred. No. 4.9e-170;  
Matches 700; Conservative 0; Mismatches 9; Indels 4; Gaps 4;  
QY 1222 GTGCTGTTCGGACACCCAGCTGCTGTCATATGTTGTAGACTTTTGATGACAAAGTACATTGTT 1281  
DB 1 GTGCTGTTCGGACACCCAGCTGCTGTCATATGTTGTAGACTTTTGATGACAAAGTACAA-TGTT 59  
QY 1282 TCTGCATCTGGGATAGAACTATAAAGGTATGAAACAAAGTACTTTGGAATTTGTAAGG 1341  
DB 60 TCTGCATCTGGGATAGAACTATAAAGGTATGAAACAAAGTACTTTGGAATTTGTAAGG 119  
QY 1342 ACCTTAAATGGACACAAACGAGGCATTGCTGTTTSCAGTACAGGACAGGCTGGTAGTG 1401  
DB 120 ACCTTAAATGGACACAAACGAGGCATTGCTGTTTSCAGTACAGGACAGGCTGGTAGTG 179  
QY 1402 AGTGGCTCATCTCAACACTATCAGATTATGGGACATAGAAATGTGTGCATGTTTACGA 1461  
DB 180 AGTGGCTCATCTCAACACTATCAGATTATGGGACATAGAAATGTGTGCATGTTTACGA 239  
QY 1462 GTGTTAGAAGCCATGAGGATTTGGTGGCTGTTGATTCGATTTGATTAACAAGAGGATGTC 1521  
DB 240 GTGTTAGAAGCCATGAGGATTTGGTGGCTGTTGATTCGATTTGATTAACAAGAGGATGTC 299  
QY 1522 AGTGGGCTCATGATGGAATAAAGTGTGGATCTTTGTGGCTGCTTTGGACCCCGT 1581  
DB 300 AGTGGGCTCATGATGGAATAAAGTGTGGATCTTTGTGGCTGCTTTGGACCCCGT 359  
QY 1582 GCTCCTCGAGGACACTCTCTACGGACCTTGTGGACATTCGGAAAGATTTTGA 1641  
DB 360 GCTCCTCGAGGACACTCTCTACGGACCTTGTGGACATTCGGAAAGATTTTGA 419  
QY 1642 CTACAGTTTGTGATGAATTCAGATTGTTCAGTAGTTTCATGATGACACAAATCTCATCTGG 1701  
DB 420 CTACAGTTTGTGATGAATTCAGATTGTTCAGTAGTTTCATGATGACACAAATCTCATCTGG 479  
QY 1702 GACTTCTTAATGATCCAGTGCCTGCAAGCTGAACCCCGCTTCCCTTCTCGAATAC 1761  
DB 480 GACTTCTTAATGATCCAGTGCCTGCAAGCTGAACCCCGCTTCCCTTCTCGAATAC 538  
QY 1762 ACCTACATCTCCAGATAAATAACCATACACTGACCTCATCTATTGTCAGGACCCATTAAA 1821  
DB 539 ACCTACATCTCCAGATAAATAACCATACACTGACCTCATCTATTGTCAGGACCCATTAAA 598  
QY 1822 GTTCGGTATTTAACGTATCTGCCAATACAGATGAGCAACAAAGTAACTCAACT 1881  
DB 599 GTTCGGTATTTAACGTATCTGCCAATACAGATGAGCAACAAAGTAACTCAACT 656  
QY 1882 ACTGCCAGTTTCCCTGGACTAGCCAGGAGGCGTTTGAGACTCTCTGTTG 1934  
DB 657 AGTGGCCAGTTTCCCTGGACTAGCCAGGAGGCGTTTGAGACTCTCTGTTG 709
```

```
RESULT 7  
CB244526  
LOCUS  
DEFINITION  
UI-M-FYO-cdg-k-22-0-UT, ri NIH_BMAP_FYO Mus musculus cDNA clone  
IMAGE: 683335 5', mRNA sequence.  
CB244526  
ACCESSION  
CB244526.1 GI:28366170
```

KEYWORDS	EST.	Mus musculus (house mouse)
SOURCE	Mus musculus	
ORGANISM	Mus musculus	
REFERENCE	1	(bases 1 to 776)
AUTHORS	NIH-MGC	http://mgi.nci.nih.gov/
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
COMMENT	Unpublished	
	Contact: Robert Strausberg, Ph.D.	
	Email: cgapbs-remail.nih.gov	
	Tissue Procurement: Dr. Jim Lin, University of Iowa	
	cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa	
	cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa	
	DNA Sequencing by: Dr. M. Bento Soares, University of Iowa	
	Clone Distribution: MGC clone distribution information can be	
	found through the I.M.A.G.E. Consortium/LLNL at:	
	http://image.llnl.gov	
	This clone was contributed by the Brain Molecular Anatomy Project	
	(BMAP)	
Seq primer: pyx-5.		
Location/Qualifiers		
1. 776		
/organism="Mus musculus"		
/mol_type="mRNA"		
/strain="C57BL/6"		
/db_xref="taxon:10090"		
/clones="IMAGE: 683135"		
/tissue_type="whole brain"		
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"		
/lab_host="DH10B (TI phage resistant)"		
/clone_lib="NIH BMAP PYO"		
/note="Organ: Brain; Vector: pyx-Asc; Site: 1: EcoR I; Site 2: Not I; The library was constructed according		
Bonaldi, Lemon and Soares, Genome Research, 6:791-806,		
1996. Denatured RNA was size fractionated on a 1% agarose		
gel. First strand cDNA synthesis was primed with oligo-dT		
primer containing a Not I site. Double strand cDNA was		
size selected according to mRNA size fraction, ligated		
with EcoR I adaptor, digested with Not I and then cloned		
directionally into pyx-Asc vector. The library tag		
sequence located between the Not I site and the polyA tail		
is AGCAGACAG. This library was created for the University		
Iowa Brain Anatomy Project (BMAP): 'Gene discovery in the		
Developing Mouse Nervous System', supported by National		
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,		
program coordinator."		
BASE COUNT	201 a	185 c 212 g 175 t 3 others
ORIGIN		
Query Match	29.9%	Score 643.4; DB 14; Length 776;
Best Local Similarity	89.2%	Pred. No. 5.2e-168;
Matches 692;	Conservative 0;	Mismatches 84; Indels 0; Gaps 0;
QY	521	GGGATTGGATCATATCGCTGAGAACATTTCTGTCATCTGATGCGATGCGAATCATCTATGTTG 580
Db	1	GGGGTCTGGACCATCGCTGAGAACATTTCTGTCATCTGATGCGATGCGAATCATCTATGTTG 60
QY	581	CTGCTGAACTTGTGTGCAAGGAATGGTACCGATGACCTCTGATGCGATGCTGTGGAAGA 640
Db	61	CTGCTGACCTGTGTGCAAGGAATGGTACCGATGACCTCTGATGCGATGCTGTGGAAGA 120
QY	641	AGCTTATCGAGGAATGGTTCAGGACAGATTTCTTGTGGAGAGCGCTTGGCAGAACGAG 700
Db	121	AGCTCATCGAGGAATGGTTCAGGACGAGCTTCTGTGGGAGCGCTTGGCAGAGCGCAG 180
QY	701	GATGGGACAGTATTATTTCAAAAACAACTCTGACGGGAATGCTCTCCCACTCTT 760
Db	181	GCTGGGACAGTATTATTTCAAAAACAACTCTGATGAGACGCTCTCCCACTCTT 240
2y	761	TATATAGACACTTTATCTTAAATTTATACAGACATTTGAGCAATAGAAATCTAATTTGA 820
b	241	TTTATAGACGCTTTATCTTAAATTTATACAGACATTTGAGCAATAGAAATTTGA 300

/notes="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dr primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is GTGCGTGGAA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BNAP). "Gene Discovery in the Developing Mouse Nervous System", supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 223 a 181 c 204 g 169 t 1 others
 ORIGIN
 Query Match 29.7%; Score 638; DB 13; Length 778;
 Best Local Similarity 89.5%; Pred. No. 1.7e-166; Indels 1; Gaps 1;
 Matches 697; Conservative 0; Mismatches 81;
 256 GCTATGAAGACTGAGAAATTTGTGGCCAAACAACTTGGCAATGGCACTTCAGTATG 315
 1 GCTATGAAGACTGAAATTTGTGGCCAAAGCCAACTTGGCAATGGCACTTCAGCATG 60
 316 ATTGTGCCAAGCAACGGAACCTTCAGCAAGCTATGAAAGGAAAGGAAGCACTGTGTGTC 375
 61 ATTGTGCCAAGCAACGCGGAACCTTCAGCAAGCTATGAGAGGAAGAGCACTGTGTGTC 120
 376 AATATCTTTGAGCAGTGGTTCAGAGTCAGATCAAGTGGAAATTTGTGGAACTTTTATATCC 435
 121 AAGTATTTTGGCAGTGGTTCAGAGTCTGATCAAGTGGAAATTTGTAGAACACCTTATATCC 180
 436 CAAATGTCTCATTTACCAATGGGCACATATACCTCGTATCTTAACTATCTTGCACAGA 495
 181 CAAATGTCTCATTTACCAATGGGCACATATACCTCGTATCTTAACTATCTTGCACAGG 240
 496 GATTTCAATCTCTCTGCCAGCTGGGGATTTGGATCATATCGTGTGAAACATTTCTGTCA 555
 241 GATTTCAATCTCTCTGCCAGCTGGGGATTTGGATCATATCGTGTGAAACATTTCTGTCA 300
 556 TACTGTGATGCCAAATCAGTATGTCTGTCTGCACTTGTCTGCAAGGATGTATCCAGGTG 615
 301 TACTGTGATGCCAAATCAGTATGTCTGTCTGCACTTGTCTGCAAGGATGTATCCAGGTG 360
 616 ACCTCTGATGGCATGCTGTGGAAGAGCTTATCGAGAGATGGTCAGGACAGATTTCTCTG 675
 361 ACCTCTGATGGCATGCTGTGGAAGAGCTTATCGAGAGATGGTCAGGACAGATTTCTCTG 420
 676 TGGAGAGGCTTGCAGAACGAGAGGATGGGACAGATTTATTTCAAAACAAACCTTCTCT 735
 421 TGGAGAGGCTTGCAGAACGAGAGGATGGGACAGATTTATTTCAAAACAAACCTTCTCT 480
 736 GAGCGGAATGCTCTCTCCCACTCTTTTATAGAGCACTTTATCTCTAAATTTATACAGAC 795
 481 GATGAGAAAGCTCTCTCCCACTCTTTTATAGAGCACTTTATCTCTAAATTTATACAGAC 540
 796 ATTGAGACATAGATCTTAATTTGAGATGTGAAGACATAGTTTACAGAGATTTACATGTC 855
 541 ATTGAGACATAGATGTGAAGATGTGAAGATGTGAAGATGTGAAGATTTACAGAGATTTCCATGTC 600
 856 CGAAGTGAACAAAGCAAGAGGTTTACTGTTTACAGATATGATGATCGAAGAAATAGTAAAGC 915
 601 CGAAGTGAACAAAGTAAAGGGTTTACTGTTTACAGATACGACGACCAAGATAGTACAGC 660
 916 GGCCTTCGACAAACACATCAGATCTGGGATTAACACATGTTGGATGTGAACGCGAATT 975
 661 GGCCTTCGAGAC-ACACCATCAAGATCTGGGANTAAAGCACACTGGAATGGAACGCGAATT 719
 976 CTCACAGGCCATACAGGTTTCAGTCTCTCTCTCCAGTATGATGAGAGAGTGTATCATAC 1034
 720 CTCACAGGCCATACAGGTTTCAGTCTCTCTCTCCAGTATGATGAGAGAGTGTATCATAC 778

RESULT 9
 LOCUS AL044095
 DEFINITION DKFP3434M1528_1 434 (synonym: htes3) Homo sapiens cDNA clone
 VERSION AL044095
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 652)
 Bloeker,H., Bocher,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
 EST (Bloeker, et al.)
 Unpublished
 Contact: Bloeker H
 MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.
 No sl sequence available.
 This clone (DKFP3434M1528) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
 Location/Qualifiers
 1..652
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFP3434M1528"
 /tissue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="434 (synonym: htes3)"
 /notes="vector: pSport1; Site_1: NotI; Site_2: SalI"
 BASE COUNT 165 a 166 c 138 g 183 t
 ORIGIN
 Query Match 29.5%; Score 635.2; DB 9; Length 652;
 Best Local Similarity 99.2%; Pred. No. 9.3e-166;
 Matches 649; Conservative 0; Mismatches 3; Indels 2; Gaps 1;
 QY 1479 GGAATTCGTGGTGTGATTCGATTTGATTAACAGAGATAGTCAGTGGGCGCTATGATGG 1538
 DB 1 GGAATTCGTGGTGTGATTCGATTTGATTAACAGAGATAGTCAGTGGGCGCTATGATGG 60
 QY 1539 AAAAATTAAGTGTGGATCTTGTGGCTGTTTGGACCCCGCTGCTCTGAGGAGCACT 1598
 DB 61 AAAAATTAAGTGTGGATCTTGTGGCTGTTTGGACCCCGCTGCTCTGAGGAGCACT 120
 QY 1599 CTGTCTACGACCTTGTGGAGCATTCGGAGAGTTTTTCGACTACAGTTTGATGATTT 1658
 DB 121 CTGTCTACGACCTTGTGGAGCATTCGGAGAGTTTTTCGACTACAGTTTGATGATTT 180
 QY 1659 CCAGATTGTTCAGTAGTTTCAATGATGACAAATCTCTATCTGGGACTTCTTAATGATCC 1718
 DB 181 CCAGATTGTTCAGTAGTTTCAATGATGACAAATC--CATCTGGACTTCTTAATGATCC 238
 QY 1719 AGCTGCCAGCTGAACCCCGCTTCCCTTCTCGAATACACCTACATCTCCAGATA 1778
 DB 239 AGCTGCCAGCTGAACCCCGCTTCCCTTCTCGAATACACCTACATCTCCAGATA 298
 QY 1779 AATAACCATACACTGACCTCATCTTGTCCCGAGACCCATTAAGTTGCGGTATTTAACT 1838
 DB 299 AATAACCATACACTGACCTCATCTTGTCCCGAGACCCATTAAGTTGCGGTATTTAACT 358

```

1839 ATCTGCCAATACAGGATGAGCAACAGTAAACAATCAAACTACTGCCAGTTTCCTGT 1898
1840 |||||||
359 ATCTGCCAATACAGGATGAGCAACAGTAAACAATCAAACTACTGCCAGTTTCCTGT 418
1899 GACTAGCCGAGGAGCAGGCTTTGAGACTCCTGTTGGGACACAGTTGCTGCGAGTCGGC 1958
419 GACTAGCCGAGGAGCAGGCTTTGAGACTCCTGTTGGGACACAGTTGCTGCGAGTCGGC 478
1959 CCAGGACGCTTACTCAGCACAACTGACTGCTTCAGTGTGCTATCAGAAGATGCTTCT 2018
479 CCAGGACGCTTACTCAGCACAACTGACTGCTTCAGTGTGCTATCAGAAGATGCTTCT 538
2019 ATCAATGTGAATGATGAGACTTTTAAACCTCCCTCTCCTCCCTTTCACCTCTGCA 2078
539 ATCTTTGTGAATGATGAGACTTTTAAACCTCCCTCTCCTCCCTTTCACCTCTGCA 598
2079 CCTAGTTTTTCCATGTTCCAGACAAAGTGACTTATAATATATATTTAGTG 2132
599 CCTAGTTTTTCCATGTTCCAGACAAAGTGACTTATAATATATATTTAGTG 652

RESULT 10
BQ425358
LOCUS BQ425358
DEFINITION AGENCOURT_7907367 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6154838
5', mRNA sequence.
ACCESSION BQ425358.1 GI:21120673
VERSION BQ425358.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 851)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13496 row: 1 column: 15
High quality sequence stop: 697.

FEATURES
source
Location/Qualifiers
1..851
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6154838"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 67"
/notes="Organ: eye; Vector: pCMV-Sport6; Site 1: Not I;
Site 2: Sal I; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
BASE COUNT 243 a 196 c 223 g 188 t 1 others
ORIGIN
Query Match 29.1%; Score 625; DB 13; Length 851;
Best Local Similarity 99.1%; Pred. No. 7.6e-163;
Matches 639; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

112 TTTATGAATTCCTCAGACAGAGAGACTGTAAATATGGCAACCCCTAGGAAGATAATA 171
172 |||||||
207 TTCCAGAATTCCTCAGACAGAGAGACTGTAAATATGGCAACCCCTAGGAAGATAATA 266

```

```

172 CCAGAGAAAGAAATTCACCTTAGACAGACATACAACAGCTGTGCCAGACTCTGCTTAAACCAA 231
173 |||||||
267 CCAGAGAAAGAAATTCACCTTAGACAGACATACAACAGCTGTGCCAGACTCTGCTTAAACCAA 326
268 |||||||
232 GAAACAGTATGTTTAGCAAGCACTGTATGAAGACTGAGAAATTTGTGTGCGCAAAACAAA 291
292 |||||||
327 GAAACAGTATGTTTAGCAAGCACTGTATGAAGACTGAGAAATTTGTGTGCGCAAAACAAA 386
328 |||||||
232 CTTGCCAATGCCATTCACAGTATGATTTGCCCAAGCAACGAAATCTCTCAGCAAGCTAT 351
351 |||||||
387 CTTGCCAATGCCATTCACAGTATGATTTGCCCAAGCAACGAAATCTCTCAGCAAGCTAT 446
388 |||||||
352 GAAPAGGAAAAGGAACTGTGTCAAAATCTTTGAGCAGTGGTCAGAGTCAGATCAAGTG 411
411 |||||||
447 GAAAGAGAAAAGGAACTGTGTCAAAATCTTTGAGCAGTGGTCAGAGTCAGATCAAGTG 506
448 |||||||
412 GAATTTGTGGAACTCTTATATCCAAATGTGTCATTAACCAATGGGCACATAAACTCG 471
471 |||||||
507 GAATTTGTGGAACTCTTATATCCAAATGTGTCATTAACCAATGGGCACATAAACTCG 566
508 |||||||
472 TATCTTAAACCTATGTTGCAGAGAGATTTCACTAATCTGCTCCAGCTCGGGATTGGAT 531
531 |||||||
567 TATCTTAAACCTATGTTGCAGAGAGATTTCACTAATCTGCTCCAGCTCGGGATTGGAT 626
568 |||||||
532 CATATCGCTGAGAACTTCTGTATACCTGGATGCCAAATCACTATGCTGCTGAATCT 591
591 |||||||
627 CATATCGCTGAGAACTTCTGTATACCTGGATGCCAAATCACTATGCTGCTGAATCT 686
628 |||||||
592 GTGTGCAAGGAATGGTACCGAGTGACCTCTGTAGGAGTCTGTGGAGAGCTTATTCGAG 651
651 |||||||
687 GTGTGCAAGGAATGGTACCGAGTGACCTCTGTAGGAGTCTGTGGAGAGCTTATTCGAG 746
688 |||||||
652 AGAATGTCAGGACAGATTTCTGTGGAGAGCTGTGGAGAGCTGTGGAGAGCTGTGGAG 710
710 |||||||
747 AGAATGTCAGGACAGATTTCTGTGGAGAGCTGTGGAGAGCTGTGGAGAGCTGTGGAG 806
748 |||||||
711 GTATTTATTCAAAACAAACCTCTCTGAGCGGAATGCTCTCTCCCA 755
755 |||||||
807 GTATTTATTCAAAACAAACCTCTCTGAGCGGAATGCTCTCTCCCA 851
852 |||||||

RESULT 11
BQ465280
LOCUS BQ465280
DEFINITION AGENCOURT_6427650 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5502075
5', mRNA sequence.
ACCESSION BQ465280
VERSION BQ465280.1 GI:18514312
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 973)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12139 row: k column: 04
High quality sequence stop: 601.
Location/Qualifiers
1..973
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

FEATURES
source

```



```

/clone="IMAGE:5502075"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_57"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
BASE COUNT      282 a      245 g      246 t
ORIGIN
Query Match      29.0%; Score 624.8; DB 12; Length 973;
Best Local Similarity 97.2%; Pred. No. 9.3e-163;
Matches 690; Conservative 0; Mismatches 12; Indels 8; Gaps 5;
2Y 658 GTCCGACAGATTTCTCTGTGGAGAGGCTGCGACAGAGAGGATGGGACAGTATTTA 717
2b 20 GTCCGACAGATTTCTCTGTGGAGAGGCTGCGACAGAGAGGATGGGACAGTATTTA 79
2Y 718 TTCAAAAACAAACCTCTGAGCGGAATGCTCTCCCACTCTTTTATAGACACTTAT 777
2b 80 TTCAAAAACAAACCTCTGAGCGGAATGCTCTCCCACTCTTTTATAGACACTTAT 139
2Y 778 CTTAAAAATTATCAAGACATTTGAGCAATAGACTTAATTGGAGATGTGAAGACATAGT 837
2b 140 CTTAAAAATTATCAAGACATTTGAGCAATAGACTTAATTGGAGATGTGAAGACATAGT 199
2Y 838 TTACAGAGATTTCACTGCGGAAGTGAAACAGCAAGAGGTTTACTGTTTACAGTATGAT 897
2b 200 TTACAGAGATTTCACTGCGGAAGTGAAACAGCAAGAGGTTTACTGTTTACAGTATGAT 259
2Y 898 GATCGAAAATAGTAAGCGGCTTCGAGACACCAATCAAGATCTCGGATAAAAACACA 957
2b 260 GATCGAAAATAGTAAGCGGCTTCGAGACACCAATCAAGATCTCGGATAAAAACACA 319
2Y 958 TTGGATGCAAGCGATTTCTACAGGCGCATACAGGTTCACTCTCTCTCCAGTATGAT 1017
2b 320 TTGGATGCAAGCGATTTCTACAGGCGCATACAGGTTCACTCTCTCTCCAGTATGAT 379
2Y 1018 GAGAGAGTGATCATAAACAGGATTCGGATTCCAGCGTTCAGAGTGTGGGATGTAATACA 1077
2b 380 GAGAGAGTGATCATAAACAGGATTCGGATTCCAGCGTTCAGAGTGTGGGATGTAATACA 439
2Y 1078 GGTGAATGCTTAAACAGTGTGATTCACCAATTGTGAAGCAGTTCTGCACTTGGTTCAAT 1137
2b 440 GGTGAATGCTTAAACAGTGTGATTCACCAATTGTGAAGCAGTTCTGCACTTGGTTCAAT 499
2Y 1138 AATGGCATGATGGTGACCTGCTCCAAAGATCGTTCCATTTGTGTATGGGATATGGGCTCC 1197
2b 500 AATGGCATGATGGTGACCTGCTCCAAAGATCGTTCCATTTGTGTATGGGATATGGGCTCC 559
2Y 1198 CCAACTGACATTTACCTCTCGGAGGTTCTGTGCGACACC-GAGCTGCTCAATGTTGT 1256
2b 560 CCAACTGACATTTACCTCTCGGAGGTTCTGTGCGACACCAGGAGCTGCTGCAATGTTGT 619
2Y 1257 AGACTTTGATGACAGTACATTG-TTTCTGCATCTGGGG--ATAGAATCTATAAGGTATG 1313
2b 620 AGACTTTGATGACAGTACATTGTTTCTGCATCTGGGGATAGAACTATAAGGTATG 679
2Y 1314 GAACA-CAAGTACTGTGAATTTGT---AAGACCTTAAATGGACACAAA 1359
2b 680 GAACACCAAGTACTGTGAATTTTGTGAAGACCTTAAATGGACACAAA 729

RESULT 12
BUS07728
LOCUS
DEFINITION BUS07728 887 bp mRNA linear EST 12-SEP-2002
5', mRNA sequence.
AGENCOURT_10095724 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6501709
ACCESSION BUS07728
VERSION BUS07728.1 GI:22813961
KEYWORDS EST.
SOURCE Homo sapiens (human)

```

```

ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 887)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M14056 row: j column: 14
High quality sequence stop: 630.
FEATURES
Location/Qualifiers
source 1..887
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6501709"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."
BASE COUNT 218 a 239 c 200 g 228 t
ORIGIN
Query Match 28.9%; Score 622.2; DB 13; Length 887;
Best Local Similarity 98.7%; Pred. No. 4.7e-162;
Matches 627; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
2Y 1509 CAAGAGGATAGTCAGTGGGGCTTATGATGGAATAAATAAGTGTGGATCTTGTGGTGC 61
2b 2 CACGCGTCCGTCAGTGGGGCTTATGATGGAATAAATAAGTGTGGATCTTGTGGTGC 61
2Y 1569 TTTGGACCCCGTCTCTGAGGACACTCTGTACGACCCCTGTGGAGCATTCGG 1628
2b 62 TTTGGACCCCGTCTCTGAGGACACTCTGTACGACCCCTGTGGAGCATTCGG 121
2Y 1629 AAGAGTTTTTCGACTACAGTTTGATGAATTCAGATTTGTCAGTAGTTCACATGATGAC 1688
2b 122 AAGAGTTTTTCGACTACAGTTTGATGAATTCAGATTTGTCAGTAGTTCACATGATGAC 181
2Y 1689 AATCCTCATCTGGAGCTTCTTAATGATCCAGTCCGACCTGAACCCCGTTCGCC 1748
2b 182 AATCCTCATCTGGAGCTTCTTAATGATCCAGTCCGACCTGAACCCCGTTCGCC 241
2Y 1749 TTCTCGAACAATACACCTTACATCTCCAGATAAATAACCATACACTGACCTCATACTTGGCC 1808
2b 242 TTCTCGAACAATACACCTTACATCTCCAGATAAATAACCATACACTGACCTCATACTTGGCC 301
2Y 1809 AGNACCCATTAAGTTGCGGTATTTAACGTATCTCCATACAGGATGAGCAACAG 1868
2b 302 AGNACCCATTAAGTTGCGGTATTTAACGTATCTCCATACAGGATGAGCAACAG 361
2Y 1869 TAACAATCAAACTACTGCGCCAGTTTCCCTGGACTAGCGAGGAGGAGGCTTTGAGACTC 1928
2b 362 TAACAATCAAACTACTGCGCCAGTTTCCCTGGACTAGCGAGGAGGAGGCTTTGAGACTC 421
2Y 1929 CTGTTGGGACACAGTTGGTCTGCACTGCGCCAGGACCGGTCTTACTCAGCACTGACTG 1988
2b 422 CTGTTGGGACACAGTTGGTCTGCACTGCGCCAGGACCGGTCTTACTCAGCACTGACTG 481
2Y 1989 CTTTCAGTGTCTGATCAGAAAGATGCTTCTATCAATTGTAATGATTGGAATTTTAAAC 2048
2b 482 CTTTCAGTGTCTGATCAGAAAGATGCTTCTATCTCTTTTGTGAATGATTGGAATTTTAAAC 541

```


2Y 2049 CTCCCTCTCTCCCTCTTACCTCTCTGACCTAGTCTTTTCCCATGTTTCCAGACAA 2108
 DB 542 CTCCCTCTCTCTCCCTCTTACCTCTCTGACCTAGTCTTTTCCCATGTTTCCAGACAA 601
 2Y 2109 GGTGACTTATAATATATATTAGTGTGTTTCCAGAA 2143
 DB 602 GGTGACTTATAATATATATTAGTGTGTTTCCAGAA 636

RESULT 13
 CB520936 744 bp mRNA linear EST 28-MAR-2003
 LOCUS UI-M-GH0-cel-18-0-UI.r1 NIH_BMAP_GH0 Mus musculus cDNA clone
 DEFINITION IMAGE: 6841315 5', mRNA sequence.
 ACCESSION CB520936
 VERSION CB520936.1 GI:29354291
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pyX-5, Location/Qualifiers
 1..744
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CS7BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE: 6841315"
 /tissue_type="Whole brain"
 /dev_stage="1, 5, and 15 days newborn"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_GH0"
 /note="Organ: Brain; Vector: pyX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 Bonaïdo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pyX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is CCACTGAAAT. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): "Gene Discovery in the
 Developing Mouse Nervous System", supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

BASE COUNT 197 a 184 c 202 g 160 t 1 others
 ORIGIN
 Query Match 28.7%; Score 616.6; DB 14; Length 744;
 Best Local Similarity 89.2%; Pred. No. 1.6e-160;
 Matches 664; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

2Y 501 CATACCTGCTCCAGCTCGGGATTGGATCATATCGCTCAGACACATTCGTGACACT 560
 DB 1 CATACCTGCTCCAGCAGCGGGTCTGGACCACATCGCTGAGAACATTCGTGACACT 60

QY 561 GGATGCCAAATCACTATGTGCTGCTGAACCTTGTGTGCRAGGAATGTAACCGAGTGAACCTC 620
 DB 61 GGAGCCCAAGTCTACTGTGTGCTGCTGAGCTCGTGTGCRAGGAATGTAACCGAGTGAACCTC 120
 QY 621 TGATGGCATCTCTGGAAGAAGCTTATCGAGAGAAATGGTCAGGACAGATTCTCTGTGGAG 680
 DB 121 GGACGGCATCTCTGGAAAAGCTCATCGAGAGGATGGTCAGGACGGACTCTCTGTGGCG 180
 QY 681 AGGCTGGCAAGACGAGAGGATGGGACAGTATTATTTCANAAACAACCTCTCTGAGCG 740
 DB 181 AGGCTGGCAGAGCGGAGAGGCTGGGGACAGTACTATTTCAAAAACAACCTCTCTGATGA 240
 QY 741 GAATGTCTCTCCCAACTCTTTTATAGACACTTTTATCCTAAAAATATACAGACATTGA 800
 DB 241 GACGCTCTCTCCCAACTCTTTTATAGAGCGCTTTATCCTAAAAATATACAGACATTGA 300
 QY 801 GACATAGATCTAATTTGAGATGGGAAGACATAGTTTACAGAGAAATCACTGCCGAAG 860
 DB 301 GACATAGATCTCAATTGGAGATGGGGGACATAGTTTACAGAGAAATCACTGCCGAAG 360
 QY 861 TGAATCAAGCAAAAGGAGTTTACTGTTTACAGTATGATGATCAGAAAAATAGTAAGCGGCT 920
 DB 361 TGAATCAAGTAAAGGGTTTACTGTTTACAGTACGACGACGAGAGATAGTCAGCGGCT 420
 QY 921 TCGAGACAAACATCAAGATCTGGGATTAACACATTTGGAATCGAAGCGAATTCCTAC 980
 DB 421 TCGAGACAAACATCAAGATCTGGGATTAACACATTTGGAATCGAAGCGAATTCCTAC 480
 QY 981 AGGCATATACAGGTTCTCTGCTCTCCAGTATGATGAGAGAGTGAATCAACAGGATC 1040
 DB 481 GGGCCACACGGGCTCCGCTCTGCTCTGAGTACGATGAGAGGGTGAATCAACAGGATC 540
 QY 1041 ATCGGATTCACCGGTGAGATGTAAATACAGTGAATGCTAAACACGCTTGAAT 1100
 DB 541 CTCAGATCCACCGTGTGGGATGTAAATACAGTGAATGCTAAACACATTTGAT 600
 QY 1101 TCACCATTTGAAAGCAGTTCTGCACTTGGGTTTCAATATGATGATGGTGGTCACTGCTC 1160
 DB 601 TCACCATTTGAAAGCAGTTCTGCACTTGGGTTTCAATATGATGATGGTGGTCACTGCTC 660
 QY 1161 CAAGATCGTTCATGCTGTATGGGATATGGCTCCCACTGACATTAACCTCCGAG 1220
 DB 661 CAAGATCGTTCATGCTGTATGGGATATGGCTCCCACTGACATTAACCTCCGAG 720
 QY 1221 GGTGCTGTGGACACCGAGCTGC 1244
 DB 721 GGTGCTGTGGACACCGAGCTGC 744

BE296484 758 bp mRNA linear EST 20-JUL-2000
 601174756F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:3530341 5',
 mRNA sequence.
 ACCESSION BE296484
 VERSION BE296484.1 GI:9180046
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 758)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LCM202 row: o column: 14
High quality sequence stop: 617.

FEATURES

Source

1..758
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3530341"
/tissue_type="thabdomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 17"
/notes="Organ: muscle; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 212 a 149 C 202 G 195 t

Query Match 28.5%; Score 613.2; DB 10; Length 758;
Best Local Similarity 96.3%; Pred. No. 1.4e-159;
Matches 649; Conservative 0; Mismatches 23; Indels 2; Gaps 2;

2y 602 AATGGTACCGAGTGACCTCTGATGGCATCTGTGGAAGAGCTTATCGAGAGATGGTCA 661
Db 1 AATGGTACCGAGTGACCTCTGATGGCATCTGTGGAAGAGCTTATCGAGAGATGGTCA 60
2y 662 GGACAGATTTCTGTGGAGAGCGCTGGGAGACGGAAGAGATGGGACAGATTATTCA 721
Db 61 GGACAGATTTCTGTGGAGAGCGCTGGGAGACGGAAGAGATGGGACAGATTATTCA 120
2y 722 AAAACAAACCTCTGACGGGAATGCTCCCAACTCTTTTATAGACACTTTATCCTA 781
Db 121 AAAACAAACCTCTGACGGGAATGCTCCCAACTCTTTTATAGACACTTTATCCTA 180
2y 782 AAATTATACAGACATTTGAGACAATAGAACTTAATTGGAGATGTGGAGACATAGTTTAC 841
Db 181 AAATTATACAGACATTTGAGACAATAGAACTTAATTGGAGATGTGGAGACATAGTTTAC 240
2y 842 AGAGATTTCACTCCCAAGTGAACAGCAAGAGAGTTTACTGTTACAGTATGATGATC 901
Db 241 AGAGATTTCACTCCCAAGTGAACAGCAAGAGAGTTTACTGTTTACAGTATGATGATC 300
2y 902 AGAAATAGTAAAGCGGCTTCGAGACACACAACTCAAGATCTGGGATAAACACATTTGG 961
Db 301 AGAAATAGTAAAGCGGCTTCGAGACACACAACTCAAGATCTGGGATAAACACATTTGG 360
2y 962 AATGCAAGCAATTTCTCAGAGCCATACAGGTTTCAGTCTGTCTCCAGTATGATGAGA 1021
Db 361 AATGCAAGCAATTTCTCAGAGCCATACAGGTTTCAGTCTGTCTCCAGTATGATGAGA 420
2y 1022 GAGTGATCAATACAGGATCATCGATTCCACGGTCAGAGTGTGGGATG-AAAACAGGT 1080
Db 421 GAGTGATCAATACAGGATCATCGATTCCACGGTCAGAGTGTGGGATG-AAAACAGGT 480
2y 1081 GAAATGCTAAACAGTTGATTACCAATTGTGAGCAGTTCTGCACCTTCGCTTCAATAAT 1140
Db 481 GAAATGCTAAACAGTTGATTACCAATTGTGAGCAGTTCTGCACCTTCGCTTCAATAAT 540
2y 1141 GGCATGATGTGACCTTGCTCCAAAGATCGTTCCATTGTGTATGGGATATGGCCTGCCCA 1200
Db 541 GGCATGATGTGACCTTGCTCCAAAGATCGTTCCATTGTGTATGGGATATGGCCTGCCCA 600
2y 1201 ACTGACATTTACCTTCGGAGGGTGTGGTC-GGACACGAGCTGCTGTCATGTTGTAGA 1259
Db 601 CTGACATTTACCTTCGGAGGGTGTGGTCGGGACCCACGAGGCTCTGTCATGTTGTGA 660
2y 1260 CTTTGTATGACAAGT 1273
Db 661 CTTTGTGTCGGGT 674

RESULT 15

LOCUS

DEFINITION

BE269297 692 bp mRNA linear EST 13-JUL-2000
601186356F1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:3544549 5',
mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Email: CGapbs-remail.nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LINL at: image.llnl.gov

Plate: LCM239 row: o column: 14

High quality sequence stop: 610.

FEATURES

source

1..692
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3544549"
/tissue_type="Burkitt lymphoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 8"
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 202 a 134 C 179 G 177 t

Query Match 28.38; Score 608.8; DB 10; Length 692;
Best Local Similarity 97.7%; Pred. No. 2.3e-158;
Matches 649; Conservative 0; Mismatches 12; Indels 3; Gaps 3;

Qy 602 AATGGTACCGAGTGACCTCTGATGGCATCTGTGGAAGAGCTTATCGAGAGATGGTCA 661
Db 1 AATGGTACCGAGTGACCTCTGATGGCATCTGTGGAAGAGCTTATCGAGAGATGGTCA 60
Qy 662 GGACAGATTTCTGTGGAGAGCGCTGGGAGACGGAAGAGATGGGACAGATTATTCA 721
Db 61 GGACAGATTTCTGTGGAGAGCGCTGGGAGACGGAAGAGATGGGACAGATTATTCA 120
Qy 722 AAAACAAACCTCTGACGGGAATGCTCCCAACTCTTTTATAGACACTTTATCCTA 781
Db 121 AAAACAAACCTCTGACGGGAATGCTCCCAACTCTTTTATAGACACTTTATCCTA 180
Qy 782 AAATTATACAGACATTTGAGACAATAGAACTTAATTGGAGATGTGGAGACATAGTTTAC 841
Db 181 AAATTATACAGACATTTGAGACAATAGAACTTAATTGGAGATGTGGAGACATAGTTTAC 240
Qy 842 AGAGATTTCACTCCCAAGTGAACAGCAAGAGAGTTTACTGTTTACAGTATGATGATC 901
Db 241 AGAGATTTCACTCCCAAGTGAACAGCAAGAGAGTTTACTGTTTACAGTATGATGATC 300
Qy 902 AGAAATAGTAAAGCGGCTTCGAGACACACAACTCAAGATCTGGGATAAACACATTTGG 961

Db	301	AGAAATAGTAAGCGCCTTCGAGACACACATCAAGATCTGGATAAAACACATGG	360
Qy	962	AATGCAAGCGAATCTCAAGGCCATACAGGTTCAAGTCTCTCTGTCTCCAGTATGATGAGA	1021
Db	361	AATGCAAGCGAATCTCAAGGCCATACAGGTTCAAGTCTCTCTGTCTCCAGTATGATGAGA	420
Qy	1022	GAGTGATCATACAGGATCATCGGATTCACCGTTCAGAGTGGGATGTAATAACAGGTG	1081
Db	421	GAGTGATCATACAGGATCATCGGATTCACCGTTCAGAGTGGGATGTAATAACAGGTG	480
Qy	1082	AAATGCTAAACAGGTTGATTCACCATTTGAAGCAGTTCTGCACCTTGGCTTCAATAATG	1141
Db	481	AAATGCTAAACAGGTTGATTCACCATTTGAAGCAGTTCTGCACCTTGGCTTCAATAATG	540
Qy	1142	GCATGATGGTGACCT-GCTCCAAAGATCGTTCATTGCTGTATGGATATGGCCTCCCCA	1200
Db	541	GCATGATGGTGACCTGGCTCCAAAGATCGTTCATTGCTGTATGGATATGGCCTCCCCA	599
Qy	1201	ACTGACATTACCTCCGAGGGTGTGTCGGACACCGAGCTGCTGTCAAATGTTGTAGAC	1260
Db	600	CTGAACATTACCTCCGAGGGTGTGTCGGACACCGAGCTGCTGTCAAATGTTGTAGAC	658
Qy	1261	TTTG	1264
Db	659	TTGG	662

Search completed: October 22, 2003, 21:57:19
 Job time : 4102 secs


```

1501 TTTGATACAGAGGATAGTCAGTGGGGCCCTATGATGGAAAAATTAAGTGTGGATCTT 1560
1501 TTTGATACAGAGGATAGTCAGTGGGGCCCTATGATGGAAAAATTAAGTGTGGATCTT 1560
1561 GTGGCTGCTTTGGGACCCCGCTCTCTGAGGAGACACTCTGTCTAGGACCCCTTTGGAG 1620
1561 GTGGCTGCTTTGGGACCCCGCTCTCTGAGGAGACACTCTGTCTAGGACCCCTTTGGAG 1620
1621 CATTCGGGAAGAGTTTTCGACTACAGTTGATGAATTCAGATTGTCTAGTATTCACAT 1680
1621 CATTCGGGAAGAGTTTTCGACTACAGTTGATGAATTCAGATTGTCTAGTATTCACAT 1680
1681 GATGACAACTCTCATCTGGGACTTCTCTAAATGATCCAGCTGCCAAAGTGAACCCCCC 1740
1681 GATGACAACTCTCATCTGGGACTTCTCTAAATGATCCAGCTGCCAAAGTGAACCCCCC 1740
1741 CGTTCCTCTCTCGAACAATACACCTACATCTCCAGATTAATACCATACATGACCTCAT 1800
1741 CGTTCCTCTCTCGAACAATACACCTACATCTCCAGATTAATACCATACATGACCTCAT 1800
1801 ACTTCCCGAGGACCCATTAAGTTGCGGTATTTAAACGATCTGCCAATACCCAGATGAGC 1860
1801 ACTTCCCGAGGACCCATTAAGTTGCGGTATTTAAACGATCTGCCAATACCCAGATGAGC 1860
1861 AACACAGTAACAATCAAACTACTGCCAGTTTCCCTGGACTAGCCGAGGAGGCTT 1920
1861 AACACAGTAACAATCAAACTACTGCCAGTTTCCCTGGACTAGCCGAGGAGGCTT 1920
1921 TGAGACTCTCTGGGACACAGTTGGTCTGAGTCGCGCCAGGACGCTTACTCAGACACA 1980
1921 TGAGACTCTCTGGGACACAGTTGGTCTGAGTCGCGCCAGGACGCTTACTCAGACACA 1980
1981 ACTGACTCTCTGAGTCTGCTATCAGAGATGCTTCTATCAATTTGTGAATGATGGAAC 2040
1981 ACTGACTCTCTGAGTCTGCTATCAGAGATGCTTCTATCAATTTGTGAATGATGGAAC 2040
2041 TTTTAAACCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100
2041 TTTTAAACCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100
2101 CAGACAAAGTGACTTATAAATATATTAGTGTGTTTCCGAGAAAAA 2151
2101 CAGACAAAGTGACTTATAAATATATTAGTGTGTTTCCGAGAAAAA 2151

RESULT 2
LAA51229
ID AAA51229 standard; DNA; 2151 bp.
AC AAA51229;
CX
CX 26-SEP-2000 (first entry)
CX Human beta-TrCP coding sequence.
CX E3 ubiquitin ligase; beta-TrCP; F-box; WD protein; I-kappa-B; inhibitor;
CX nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;
CX anti-inflammatory; immunosuppressive; cytostatic; ds.
CX Homo sapiens.
CX
CX Key Location/Qualifiers
CX CDS 70..1779
CX /tag= a
CX /product= Human_TrCP
CX
CX WO20003447-A2.
CX
CX 15-JUN-2000.
CX
CX 10-DEC-1999; 99WO-US29371.
CX
CX 10-DEC-1998; 98US-0210060.

```

```

XX
PA (SIGN-) SIGNAL PHARM INC.
PA (VISS) VISSUM RES & DEV CO.
XX
PI Manning AM, Mercurio F, Amit S, Ben-eriah Y, Davis M, Hatzubai A;
PI Lavon I, Yaron A;
XX
DR WPI; 2000-431294/37.
DR P-ESDB; AAY96697.
XX
PT Polypeptide enhancing phosphorylated I-kappaB ubiquitination useful for
PT treating disorder associated with NF-kappaB activation e.g. cancer,
PT comprising amino acid sequence of human E3 ubiquitin ligase or its
PT variant
XX
PS Disclosure; Page 72; 77pp; English.
XX
CC This DNA encodes human beta-TrCP, an F-box/WD protein family member,
CC which has been shown to have homology to human E3 ubiquitin ligase (E3).
CC E3 enhances ubiquitination of phosphorylated I-kappa-B, an inhibitor
CC protein of nuclear factor kappa-B (NF-kappa-B). Understanding I-kappa-B
CC degradation via the ubiquitin pathway is useful for identifying
CC modulators of this process for use in treating diseases associated with
CC activation of NF-kappa-B. In vitro analysis suggests that deletion of
CC the F-box results in a protein that functions as a dominant negative
CC molecule in vivo. Transient over-expression of endogenous I-kappa-B-alpha
CC in stimulated Jurkat cells, resulting in accumulation of phosphorylated
CC I-kappa-B-alpha. E3 can be used to screen for modulators of NF-kappa-B
CC activity. E3 and beta-TrCP can be used to modulate NF-kappa-B to treat
CC inflammatory diseases, autoimmune diseases, cancer and viral infections.
XX
SQ Sequence 2151 bp; 628 A; 467 C; 513 G; 543 T; 0 other;
XX
Query Match 100.0%; Score 2151; DB 21; Length 2151;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGCGTTGGCTGGCGCTTGGCACCACCAAGGGGGCGGCGGCGGAGAGCGGACCCAGTGCC 60
DB 1 TGCGTTGGCTGGCGCTTGGCACCACCAAGGGGGCGGCGGCGGAGAGCGGACCCAGTGCC 60
QY 61 TCGGCGATTATGGACCGCGCGGCGGCGGCGGCGGCGGAGAGCGGACCCAGTGAT 120
DB 61 TCGGCGATTATGGACCGCGCGGCGGCGGCGGCGGCGGAGAGCGGACCCAGTGAT 120
QY 121 TCCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
DB 121 TCCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 181 AATTCACCTTAGACAGACATACACAGCTGTGCCAGACTCTGCTTTAAACCAAGAAACAGTA 240
DB 181 AATTCACCTTAGACAGACATACACAGCTGTGCCAGACTCTGCTTTAAACCAAGAAACAGTA 240
QY 241 TGTTCAGACAGCACTGCTATGAAGACTGAGAGATTGTTGGCCCAACAAACAACTTCCCAAT 300
DB 241 TGTTCAGACAGCACTGCTATGAAGACTGAGAGATTGTTGGCCCAACAAACAACTTCCCAAT 300
QY 301 GGCACCTTCCAGTATGATTGTGCCCAAGCAACGAAACCTCTCAGCAAGCTATGAAAGGAA 360
DB 301 GGCACCTTCCAGTATGATTGTGCCCAAGCAACGAAACCTCTCAGCAAGCTATGAAAGGAA 360
QY 361 AAGGAACTGTCTCAAAATACATTTTGGAGAGTGTGTGAGAGAGAGAGAGAGAGAGAGAG 420
DB 361 AAGGAACTGTGTCAAAATACATTTTGGAGAGTGTGTGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 GAACATCTTATATCCCAAAATGTCTATTACCAACATGGGCAATATATCTCGTATCTTAAA 480
DB 421 GAACATCTTATATCCCAAAATGTCTATTACCAACATGGGCAATATATCTCGTATCTTAAA 480
QY 481 CCTATGTTGCAGAGAGATTTCATACCTGTCTGCCAGCTCGGGGAGTGGATCATATCGCT 540
DB 481 CCTATGTTGCAGAGAGATTTCATACCTGTCTGCCAGCTCGGGGAGTGGATCATATCGCT 540

```

2y	541	GAGAACATTCGTGTCATACCTGGATGCCAAATCACTATGTGCTGCTGAACCTGTGTGCAAG	600
2b	541	GAGAACATTCGTGTCATACCTGGATGCCAAATCACTATGTGCTGCTGAACCTGTGTGCAAG	600
2y	601	GAATGGTACCAGTGAACCTCTGTATGGATGCTGTGGAAGAGCTTATCGAGAGAAATGGTC	660
2b	601	GAATGGTACCAGTGAACCTCTGTATGGATGCTGTGGAAGAGCTTATCGAGAGAAATGGTC	660
2y	661	AGGACAGATTCTCTGTGGAGAGCCCTGGCAGACAGAGAGGATGGGACAGTATTATATTC	720
2b	661	AGGACAGATTCTCTGTGGAGAGCCCTGGCAGACAGAGAGGATGGGACAGTATTATATTC	720
2y	721	AAAAACAAACCTCCTGACGGGAATGCTCTCCCAACTCTTTTATAGACACTTTATCCCT	780
2b	721	AAAAACAAACCTCCTGACGGGAATGCTCTCCCAACTCTTTTATAGACACTTTATCCCT	780
2y	781	AAATTTATAGAACACATTGAGACAATAGAAATCTAATTGGAGATGTGGAGACATAGTTTA	840
2b	781	AAATTTATAGAACACATTGAGACAATAGAAATCTAATTGGAGATGTGGAGACATAGTTTA	840
2y	841	CAGAGAAATTCACCTCCGAAGTGAACAAAGCAAGAGGATTTACTGTTTACAGTATGATGAT	900
2b	841	CAGAGAAATTCACCTCCGAAGTGAACAAAGCAAGAGGATTTACTGTTTACAGTATGATGAT	900
2y	901	CAGAAATAGTAACGGGCCCTTCAGACACAACATCAAGATCTGGGATAAAAACAATG	960
2b	901	CAGAAATAGTAACGGGCCCTTCAGACACAACATCAAGATCTGGGATAAAAACAATG	960
2y	961	GAATGCAAGCAATTCCTCAGAGCCCATACAGGTTTCAGTCCCTCTGTCTCCAGTATGATGAG	1020
2b	961	GAATGCAAGCAATTCCTCAGAGCCCATACAGGTTTCAGTCCCTCTGTCTCCAGTATGATGAG	1020
2y	1021	AGAGTGATCAACAGAGATCATCGGATCCACGGTCCAGAGTGGGATGTAAATACAGGT	1080
2b	1021	AGAGTGATCAACAGAGATCATCGGATCCACGGTCCAGAGTGGGATGTAAATACAGGT	1080
2y	1081	GAATGCTAAACACGGTTGATTACCACTGTGGAAGCAGTTCTGCACTTCGGTTCAATTAAT	1140
2b	1081	GAATGCTAAACACGGTTGATTACCACTGTGGAAGCAGTTCTGCACTTCGGTTCAATTAAT	1140
2y	1141	GGCATGATGGTGACCTGCTCCAAAGATCGTCCATTGCTGATGGGATAGGCTCCCGCA	1200
2b	1141	GGCATGATGGTGACCTGCTCCAAAGATCGTCCATTGCTGATGGGATAGGCTCCCGCA	1200
2y	1201	ACTGACATTACCCCTCCGAGGGGTGCTGGTGGACACCGAGCTGCTGTCAATGTGTAGAC	1260
2b	1201	ACTGACATTACCCCTCCGAGGGGTGCTGGTGGACACCGAGCTGCTGTCAATGTGTAGAC	1260
2y	1261	TTTGATGACAAAGTACATTGTTCTGATCTGGGATAGAACTATAAGGATGTGGAACACA	1320
2b	1261	TTTGATGACAAAGTACATTGTTCTGATCTGGGATAGAACTATAAGGATGTGGAACACA	1320
2y	1321	AGTACTGTGGAATTTGTAAGGACCTTAATGACACAAACGAGGCAATGCTGTTTTCGAG	1380
2b	1321	AGTACTGTGGAATTTGTAAGGACCTTAATGACACAAACGAGGCAATGCTGTTTTCGAG	1380
2y	1381	TACAGGGACAGGCTGGTAGTCAGTGGCTCATCTGACACAACCTATCAGATTATGGGACATA	1440
2b	1381	TACAGGGACAGGCTGGTAGTCAGTGGCTCATCTGACACAACCTATCAGATTATGGGACATA	1440
2y	1441	GAATGTGTGATGTTTACAGATGTTTAGAAGGCCATGAGGAATTTGGTCTGTATTCGA	1500
2b	1441	GAATGTGTGATGTTTACAGATGTTTAGAAGGCCATGAGGAATTTGGTCTGTATTCGA	1500
2y	1501	TTTGATTAACAAGAGGATAGTCAGTGGGGCCCTATGATGGAAAAATTAAGTGTGGGATCTTT	1560
2b	1501	TTTGATTAACAAGAGGATAGTCAGTGGGGCCCTATGATGGAAAAATTAAGTGTGGGATCTTT	1560
2y	1561	GTGGCTGCTTTGGACCCCGCTCTCTGACGGGACACTCTGTCTACGAGACCTTGTGGAG	1620
2b	1561	GTGGCTGCTTTGGACCCCGCTCTCTGACGGGACACTCTGTCTACGAGACCTTGTGGAG	1620

QY	1621	CATTCCGGAAGAGTTTTTCGACTACAGTTTGATGAATTCAGATTGTCAGTAGTTCACAT	1680
DB	1621	CATTCCGGAAGAGTTTTTCGACTACAGTTTGATGAATTCAGATTGTCAGTAGTTCACAT	1680
QY	1691	GATGACACAATCTCATCTGGGACTTCTTAATGATGCAGCTGCCCAAGCTGAACCCCCC	1740
DB	1691	GATGACACAATCTCATCTGGGACTTCTTAATGATGCAGCTGCCCAAGCTGAACCCCCC	1740
QY	1741	CGTTCCCTTCTCGAACAATACACTACATCTCCAGATAAATAACATACACTGACCTCAT	1800
DB	1741	CGTTCCCTTCTCGAACAATACACTACATCTCCAGATAAATAACATACACTGACCTCAT	1800
QY	1801	ACTTGCCCCAGGACCCATTAAAGTTGGGTATTTAAAGTATCTGCAATACCCAGGATGAGC	1860
DB	1801	ACTTGCCCCAGGACCCATTAAAGTTGGGTATTTAAAGTATCTGCAATACCCAGGATGAGC	1860
QY	1861	AACACACAGTAAACATCAAACTACTGCCGAGTTCCCTGGACTAGCCGAGGAGGCGCTT	1920
DB	1861	AACACACAGTAAACATCAAACTACTGCCGAGTTCCCTGGACTAGCCGAGGAGGCGCTT	1920
QY	1921	TGAGACTCCTGTTGGGACACAGTTGGTCTGACAGTCGGGCCAGGACGGTCTACTCAGCAC	1980
DB	1921	TGAGACTCCTGTTGGGACACAGTTGGTCTGACAGTCGGGCCAGGACGGTCTACTCAGCAC	1980
QY	1981	ACTGACGCTTCAGTCTGCTGCTATCAGAAAGATCTCTTCTCAATTGTGATGATGGAAC	2040
DB	1981	ACTGACGCTTCAGTCTGCTGCTATCAGAAAGATCTCTTCTCAATTGTGATGATGGAAC	2040
QY	2041	TTTTAAACCTCCCTCCTCTCTCTCTCTTCCACCTCTGCACCTAGTTTTTCCCATTTGGTTC	2100
DB	2041	TTTTAAACCTCCCTCCTCTCTCTCTCTTCCACCTCTGCACCTAGTTTTTCCCATTTGGTTC	2100
QY	2101	CAGACAAAGGTGACTTATAAATATATTAGTGTGTTTTCGCAAGAAAAA	2151
DB	2101	CAGACAAAGGTGACTTATAAATATATTAGTGTGTTTTCGCAAGAAAAA	2151
RESULT 3			
AAZ93350			
ID	AAZ93350 standard; cDNA; 2151 BP.		
XX	AAZ93350;		
AC	AAZ93350;		
XX			
DT	16-AUG-2000 (first entry)		
DE	Sequence encoding F-box protein FBP-1.		
XX			
KW	F-box protein; FBP; diagnosis; treatment; screening; agonist;		
KW	antagonist; proliferative disorder; differentiative disorder;		
KW	breast cancer; prostate cancer; ovarian cancer; cancer;		
KW	small cell lung carcinoma; immune disorder; cardiovascular disorder;		
KW	inflammatory disorder; human; ss.		
XX			
OS	Homo sapiens.		
XX			
Key	Location/Qualifiers		
FT	70..1779		
CDS	/*tag= a		
FT	/*product= FBP-1		
FT			
XX			
XX	WC200012679-A1.		
PN			
XX			
XX	509-MAR-2000		
PD			
XX			
XX	27-AUG-1999; 99WO-US19560.		
XX			
XX	28-AUG-1998; 98US-0098355.		
PR	03-FEB-1999; 99US-0118568.		
PR	15-MAR-1999; 99US-0124449.		
XX			
XX	(UJNY) UNIV NEW YORK STATE.		
PA			
XX			
PI	Chiaur DS, Pagano M, Latres E;		

X WPI: 2000-256635/22.
R P-PSDB; AAY83041.
X

T Novel nucleic acid for screening compounds useful for treating
T proliferative and differentiative disorders such as cancer and immune
T disorders comprises sequences encoding ubiquitin ligases -
X

S Disclosure; Figure 3b, 3c, 3d; 245pp; English.

X Nucleic acids encoding substrate-targeting subunits of ubiquitin
X ligases with F-box motifs (F-box proteins) are useful for diagnosis
X of proliferative and differentiative related disorders by measuring
X FBP gene expression. Cells expressing such proteins or
X their fragments are useful for screening compounds. The compounds
X are agonists or antagonists, which are useful for treating a
X proliferative or differentiative disorder in a mammal such as
X breast, ovarian and prostate cancer and small cell lung carcinoma
X and also major opportunistic infections, immune disorders,
X cardiovascular diseases and inflammatory disorders. FBP protein,
X analogs, derivatives and their subsequences, anti-FBP antibodies
X are also useful in diagnosis of the disorders.

X Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;

Query Match 100.0%; Score 2151; DB 21; Length 2151;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 1 TCGGTGCTGGCCTGGACCAAGGGCGGCCCGGGAGAGGGAGCCAGTGGCC 60
b 1 TCGGTGCTGGCCTGGACCAAGGGCGGCCCGGGAGAGGGAGCCAGTGGCC 60
2Y 61 TCGGGATTATGGAACCGCGCGAGCGGTGCTGCAAGAGAGGCACTCAAGTTTATGAT 120
b 61 TCGGGATTATGGAACCGCGCGAGCGGTGCTGCAAGAGAGGCACTCAAGTTTATGAT 120
2Y 121 TCCTCAGAGAGAGAGACTGTAAATAGCGAAACCCCTAGGAAGATATACAGAGAG 180
b 121 TCCTCAGAGAGAGAGACTGTAAATAGCGAAACCCCTAGGAAGATATACAGAGAG 180
2Y 181 AATTCACCTTAGACAGACATACACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTA 240
b 181 AATTCACCTTAGACAGACATACACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTA 240
2Y 241 TGTTTAGCAGCAGCTGTATGAAGACTGAGAAATTTGTGCGCAAAACAAACTTGCAT 300
b 241 TGTTTAGCAGCAGCTGTATGAAGACTGAGAAATTTGTGCGCAAAACAAACTTGCAT 300
2Y 301 GGCACCTCCAGTATGATTTGCCCAAGCAAGCAAGCACTCTCAGCAAGCTATGAAGAGAA 360
b 301 GGCACCTCCAGTATGATTTGCCCAAGCAAGCAAGCACTCTCAGCAAGCTATGAAGAGAA 360
2Y 361 AAGGAACCTGTGTGTCAAATATTTTGGCAGTGTGTCAGAGTCAAGTGGAAATTTGTG 420
b 361 AAGGAACCTGTGTGTCAAATATTTTGGCAGTGTGTCAGAGTCAAGTGGAAATTTGTG 420
2Y 421 GAACATCTTATATCCCAATCTGTCATACCAATGCGGACATACATACATACATCTTAA 480
b 421 GAACATCTTATATCCCAATCTGTCATACCAATGCGGACATACATACATCTTAA 480
2Y 481 CTTATGTTGCGAGAGATTTTCAATACCTGTGCGAGCTCGGGAGTTGGATCATATCGCT 540
b 481 CTTATGTTGCGAGAGATTTTCAATACCTGTGCGAGCTCGGGAGTTGGATCATATCGCT 540
2Y 541 GAGACATCTGTACATCTGATGCGCAATATCATATGCTGCTGCTGCTGCTGCTGCTGCTG 600
b 541 GAGACATCTGTACATCTGATGCGCAATATCATATGCTGCTGCTGCTGCTGCTGCTGCTG 600
2Y 601 GAATGGTACCGAGTACCTCTGATGGGATGCTGTTGGAGAGAGCTTATCGAGAGAGTGGTC 660
b 601 GAATGGTACCGAGTACCTCTGATGGGATGCTGTTGGAGAGAGCTTATCGAGAGAGTGGTC 660

QY 661 AGGACAGATTCTCTGTGGAGAGGCTTGGCAGAACGAGAGGATGGGACAGTATTATTTC 720
Db 661 AGGACAGATTCTCTGTGGAGAGGCTTGGCAGAACGAGAGGATGGGACAGTATTATTTC 720
QY 721 AAAAACAACCTCTGACGGGATGCTCTCCCAACTCTTTTATAGAGCACTTTATCCT 780
Db 721 AAAAACAACCTCTGACGGGATGCTCTCCCAACTCTTTTATAGAGCACTTTATCCT 780
QY 781 AAAATTATACAAGACATTGGAGACAATAGAAATCTAATTGGAGATGTGGAGACATAGTTTA 840
Db 781 AAAATTATACAAGACATTGGAGACAATAGAAATCTAATTGGAGATGTGGAGACATAGTTTA 840
QY 841 CAGAGAATTCACTGCGGAGTGAACAACAAGAGAGTTTACTGTTTACAGTATGATGAT 900
Db 841 CAGAGAATTCACTGCGGAGTGAACAACAAGAGAGTTTACTGTTTACAGTATGATGAT 900
QY 901 CAGAAAATAGTAAGCGCCTTGGAGACAACAANTCAAGATCTGGGATAAAAACACATTG 960
Db 901 CAGAAAATAGTAAGCGCCTTGGAGACAACAANTCAAGATCTGGGATAAAAACACATTG 960
QY 961 GAATGCAAGGAAATCTCAGAGCCATACAGTTTCAAGTTCAGTCTCTGCTCCAGTATGATGAG 1020
Db 961 GAATGCAAGGAAATCTCAGAGCCATACAGTTTCAAGTTCAGTCTCTGCTCCAGTATGATGAG 1020
QY 1021 AGAGTGATCATACAGGATCATCGGATTCACGGTCCAGAGTGGGATGTAATACAGGT 1080
Db 1021 AGAGTGATCATACAGGATCATCGGATTCACGGTCCAGAGTGGGATGTAATACAGGT 1080
QY 1081 GAATGCTAAACAGTTGATTCACCAATTTGAAGCAGTTCTGCACTTCGTTTCAATTAAT 1140
Db 1081 GAATGCTAAACAGTTGATTCACCAATTTGAAGCAGTTCTGCACTTCGTTTCAATTAAT 1140
QY 1141 GGCATGATGTCACCTCTCCAAAGATCGTTTCAATTTGCTGATGGGATGTCCTCCCA 1200
Db 1141 GGCATGATGTCACCTCTCCAAAGATCGTTTCAATTTGCTGATGGGATGTCCTCCCA 1200
QY 1201 ACTGACATTAACCTCCGAGAGGCTGCTGTCGAGACACCGAGCTGCTGCTCAATGTTGAGC 1260
Db 1201 ACTGACATTAACCTCCGAGAGGCTGCTGTCGAGACACCGAGCTGCTGCTCAATGTTGAGC 1260
QY 1261 TTTGATGACAAGTACATTGTTTCTGCATCTGGGATAGAACTATAAAGTATGGAACACA 1320
Db 1261 TTTGATGACAAGTACATTGTTTCTGCATCTGGGATAGAACTATAAAGTATGGAACACA 1320
QY 1321 AGTACTTTGGAATTTGTAAGGACCTTAAATGGAACAACAACGAGGCAATTCCTGTTTGCAG 1380
Db 1321 AGTACTTTGGAATTTGTAAGGACCTTAAATGGAACAACAACGAGGCAATTCCTGTTTGCAG 1380
QY 1381 TACAGGACAGGCTGTGATGAGTGGCTCATCTGACACACTATCAGATTATGGGACATA 1440
Db 1381 TACAGGACAGGCTGTGATGAGTGGCTCATCTGACACACTATCAGATTATGGGACATA 1440
QY 1441 GAATGTTGTCATGTTTACAGTGTGTAGAGCCATAGAGAAATTCGTCGTTGATTCGA 1500
Db 1441 GAATGTTGTCATGTTTACAGTGTGTAGAGCCATAGAGAAATTCGTCGTTGATTCGA 1500
QY 1501 TTTGATTAACAGAGGATGTCAGTGGGCTATGATGGAAGAAATTAAGTGTGGGATCTT 1560
Db 1501 TTTGATTAACAGAGGATGTCAGTGGGCTATGATGGAAGAAATTAAGTGTGGGATCTT 1560
QY 1561 GTGCTGCTTTGGACCCCGCTGCTCCTGCGAGGACACTCTGTCACGGACCCCTTGTGGAG 1620
Db 1561 GTGCTGCTTTGGACCCCGCTGCTCCTGCGAGGACACTCTGTCACGGACCCCTTGTGGAG 1620
QY 1621 CATTCGGAGAGTGTTCGACTACAGTTTGAATTCAGATTCAGATTCAGTATTCAT 1680
Db 1621 CATTCGGAGAGTGTTCGACTACAGTTTGAATTCAGATTCAGTATTCAT 1680
QY 1681 GATGACACATCTCTATCTGGGACTTCCTAAATGATCCAGCTGCCCAAGCTGAACCCCCC 1740
Db 1681 GATGACACATCTCTATCTGGGACTTCCTAAATGATCCAGCTGCCCAAGCTGAACCCCCC 1740
QY 1741 CGTTCCCTCTTCGAACATACACCTACATCTCCAGATAAATACCATACACTGACCTCAT 1800

1741 CGTTCCCTTCGAAACATACACCTACATCTCCAGATTAATACCATACATGACCTCAT 1800
 1801 ACTTGGCCAGGACCCATTAAAGTTGGGTAATTAACGTAATCTGCAATACAGGATGAGC 1860
 1801 ACTTGGCCAGGACCCATTAAAGTTGGGTAATTAACGTAATCTGCAATACAGGATGAGC 1860
 1861 AACACAGTAACAACTACTGCCAGTTTCCCTGGACTAGCCGAGGAGCAGGCGTT 1920
 1861 AACACAGTAACAACTACTGCCAGTTTCCCTGGACTAGCCGAGGAGCAGGCGTT 1920
 1921 TGAGACTCTCTGTGGACACAGTTGGTCTGAGTGGCCGACGAGCGTCTACTCAGACA 1980
 1921 TGAGACTCTCTGTGGACACAGTTGGTCTGAGTGGCCGACGAGCGTCTACTCAGACA 1980
 1981 ACTGACTGTCTTCTAGTCTCTATCAGAGATGCTCTTCTATCAATTTGAATGTGGAAC 2040
 1981 ACTGACTGTCTTCTAGTCTCTATCAGAGATGCTCTTCTATCAATTTGAATGTGGAAC 2040
 2041 TTTTAAACCTCCCT 2100
 2041 TTTTAAACCTCCCT 2100
 2101 CAGACAAAGTGACTTATAAATATATTTAGTGTCTTTTGCAGAAAAAAA 2151
 2101 CAGACAAAGTGACTTATAAATATATTTAGTGTCTTTTGCAGAAAAAAA 2151

RESULT 4

AAZ93710
 ID RAZ93710 standard; DNA; 2151 BP.

AAZ93710;

DT 16-AUG-2000 (first entry)

F -box protein hBetaTrCp coding sequence.

Ubiquitin ligase; SCF, F-box protein; targeted degradation;
 destabilisation; proteolysis; drug discovery; gene therapy; cancer;
 oncoprotein; Huntington's disease; gene knockout; delivery systems;
 human; ss.

Homo sapiens.

Key Location/Qualifiers
 CDS 70..1779

/*tag= a
 /product= F-box protein hBetaTrCp

WC200022110-A2.

20-APR-2000.

08-OCT-1999; 99WC-US23705.

09-OCT-1998; 98US-0103787.

(HARD) HARVARD COLLEGE.

Zhou P, Howley P;

WPI; 2000-317970/27.

P-PSDB; AAY83250.

Targeting degradation of polypeptide useful for treating cancer and
 other proliferative disorders, involves conjugating polypeptide with
 ubiquitin protein ligase or inhibiting ubiquitination using organic
 compound

Claim 10; Page 171; 185pp; English.

The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin

CC ligases) which can be used for the targetted degradation of a target
 CC polypeptide in vivo. Targetted degradation is achieved by expressing
 CC the ubiquitin ligase in a cell linked to the interaction domain of
 CC the target polypeptide and thereby recruiting the target polypeptide
 CC to the ubiquitin ligase. Such methods are useful for decreasing or
 CC increasing the level of a target polypeptide and for creating and
 CC expressing a destabilized polypeptide which is subjected to SCF
 CC mediated proteolysis. Degrading any desired protein in a cell is
 CC useful for preventing or treating diseases caused by the presence of
 CC abnormal amount of the specific polypeptides, for drug discovery and
 CC for gene therapy. Diseases treated include cancer, by degradation of
 CC oncoproteins, Huntington's disease, other proliferative disorders and
 CC microbial infections. The method provides a quick and easy
 CC alternative to gene knockout technology. The target polypeptide can
 CC be degraded at all stages, or a specific stage, of development in the
 CC mature animal.
 XX
 SQ Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;
 Query Match 100.0%; Score 2151; DB 21; Length 2151;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGCCTTGGCTGCGGCTTGGACCAAAAGGGGCGGCCCGCGGAGAGCGGACCCAGTGCC 60
 DB 1 TGCCTTGGCTGCGGCTTGGACCAAAAGGGGCGGCCCGCGGAGAGCGGACCCAGTGCC 60
 QY 61 TCGGCGATTATGACCCCGCGGAGCGGCTGCTCAAGAGAGAGGCACTCAAGTTTATGAAT 120
 DB 61 TCGGCGATTATGACCCCGCGGAGCGGCTGCTCAAGAGAGAGGCACTCAAGTTTATGAAT 120
 QY 121 TCCTCAGAGAGAGAGACTGTATTAATGGCGAACCCCTTAGGAAGATAATACAGAGAAG 180
 DB 121 TCCTCAGAGAGAGAGACTGTATTAATGGCGAACCCCTTAGGAAGATAATACAGAGAAG 180
 QY 181 AATTCACCTTAGACAGACATACACAGCTGTGCAGACTCTGCTTAAACCAAGAACAGTA 240
 DB 181 AATTCACCTTAGACAGACATACACAGCTGTGCAGACTCTGCTTAAACCAAGAACAGTA 240
 QY 241 TGTTAGCAGAGACTGTATGAAGACTGTGAATTTGGTGGCCAAACAAACCTCCCAAT 300
 DB 241 TGTTAGCAGAGACTGTATGAAGACTGTGAATTTGGTGGCCAAACAAACCTCCCAAT 300
 QY 301 GGCATTCCAGTATGATTGTGCCCAAGCAAGCAAACTCTCAGCAAGCTATGAAAAGAA 360
 DB 301 GGCATTCCAGTATGATTGTGCCCAAGCAAGCAAACTCTCAGCAAGCTATGAAAAGAA 360
 QY 361 AAGGAACCTGTGTCAAAATACCTTTGAGCGAGTGGTCAAGTCAAGTGAATTTGTG 420
 DB 361 AAGGAACCTGTGTCAAAATACCTTTGAGCGAGTGGTCAAGTCAAGTGAATTTGTG 420
 QY 421 GAACATCTTATATCCCAAAATGTCTCATTACCAACATGGGCAATAACTCGTATCTTAA 480
 DB 421 GAACATCTTATATCCCAAAATGTCTCATTACCAACATGGGCAATAACTCGTATCTTAA 480
 QY 481 CCTATGTGCAGAGAGATTTTCAATGCTCTGCCAGCTCGGGGATTTGGATCATTCCT 540
 DB 481 CCTATGTGCAGAGAGATTTTCAATGCTCTGCCAGCTCGGGGATTTGGATCATTCCT 540
 QY 541 GAGAACATTTCTGTACATCTGGATGCGCAAAATCACTATGTCTGCTGAACCTGTGTGCAAG 600
 DB 541 GAGAACATTTCTGTACATCTGGATGCGCAAAATCACTATGTCTGCTGAACCTGTGTGCAAG 600
 QY 601 GAATGGTACCGAGTACCTCTGATGGCATGCTGTGGAAGAACTTATCGAGAAATGGTC 660
 DB 601 GAATGGTACCGAGTACCTCTGATGGCATGCTGTGGAAGAACTTATCGAGAAATGGTC 660
 QY 661 AGGACAGATTTCTGTGGAGAGCGCTTGGCAGAACCGAAGAGGATGGGACAGTATTATTC 720
 DB 661 AGGACAGATTTCTGTGGAGAGCGCTTGGCAGAACCGAAGAGGATGGGACAGTATTATTC 720
 QY 721 AAAAACAACCTCTGTGACGGGAATGTCTCTCCAACTCTTTTATAGAGCACTTATCCT 780

```

1b 721 AAAAAAACCCTCTGACGGGAAGCTCCTCCCAACTCTTTTATAGAGCACTTTATCCT 780
1b 781 AAAATTATACAGACATTTGAGACAAATGAATCTAATTGGAGATGTGGAACATAGTTTA 840
1b 781 AAAATTATACAGACATTTGAGACAAATGAATCTAATTGGAGATGTGGAACATAGTTTA 840
1b 841 CAGAGAAATTCATGCGGAGTGAACAAAGCAAGAGAGTTTACTGTTTACAGTATGATGAT 900
1b 841 CAGAGAAATTCATGCGGAGTGAACAAAGCAAGAGAGTTTACTGTTTACAGTATGATGAT 900
1b 901 CAGAAAATAGTAAGGGGCTTCGACACAAACAAATCAAGATCTGGGATAAAACACATTG 960
1b 901 CAGAAAATAGTAAGGGGCTTCGACACAAACAAATCAAGATCTGGGATAAAACACATTG 960
1b 961 GAATGCAAGCAATTTCTACAGGCCATACAGGTTTCCAGGTCAGAGTGGGATGTAATACAGGT 1020
1b 961 GAATGCAAGCAATTTCTACAGGCCATACAGGTTTCCAGGTCAGAGTGGGATGTAATACAGGT 1020
1b 1021 AGAGTGATCATACAGGATCATCGGATTTCCAGGTCAGAGTGGGATGTAATACAGGT 1080
1b 1021 AGAGTGATCATACAGGATCATCGGATTTCCAGGTCAGAGTGGGATGTAATACAGGT 1080
1b 1081 GAAATGCTAAACACGTTGATTCACCAATTTGTGAAGCAGTCTGCACTTGCCTTTCAATAAT 1140
1b 1081 GAAATGCTAAACACGTTGATTCACCAATTTGTGAAGCAGTCTGCACTTGCCTTTCAATAAT 1140
1b 1141 GGCATGATGCTGCTGCTCCAAAGATCGTTCCATTTGCTGTATGGGATATGGCTCCCA 1200
1b 1141 GGCATGATGCTGCTGCTCCAAAGATCGTTCCATTTGCTGTATGGGATATGGCTCCCA 1200
1b 1201 ACTGACATTTACCTTCGGAGGCTGCTGCTGGAACACCGAGTCTGCTCAATGTTGTAGAC 1260
1b 1201 ACTGACATTTACCTTCGGAGGCTGCTGCTGGAACACCGAGTCTGCTCAATGTTGTAGAC 1260
1b 1261 TTGATGCAAGTACATTTCTGCACTTGGGATGAACTATAAAGGTATGGAACACA 1320
1b 1261 TTGATGCAAGTACATTTCTGCACTTGGGATGAACTATAAAGGTATGGAACACA 1320
1b 1321 AGTACTGTGAATTTGTAAAGCACTTAAATGGAACAAACAGGCACTTGCTGTTTCAG 1380
1b 1321 AGTACTGTGAATTTGTAAAGCACTTAAATGGAACAAACAGGCACTTGCTGTTTCAG 1380
1b 1381 TACAGGACAGGCTGCTGCTGAGTGGCTCATCTGACACACATATCAGATTTGGACATA 1440
1b 1381 TACAGGACAGGCTGCTGCTGAGTGGCTCATCTGACACACATATCAGATTTGGACATA 1440
1b 1441 GAATGTGTCATGTTTACAGTGTGTAGAGGCCATGAGGAATTTGCTGCTGTTATTCGA 1500
1b 1441 GAATGTGTCATGTTTACAGTGTGTAGAGGCCATGAGGAATTTGCTGCTGTTATTCGA 1500
1b 1501 TTTGATAACAGAGGATGCTGCTGAGTGGGCTATGATGGAATAATTAAGTGGGATCTT 1560
1b 1501 TTTGATAACAGAGGATGCTGCTGAGTGGGCTATGATGGAATAATTAAGTGGGATCTT 1560
1b 1561 GTGCTGCTTTGACCCCGCTGCTGCTGAGGACACTCTGCTACGAGCCCTTGTGAG 1620
1b 1561 GTGCTGCTTTGACCCCGCTGCTGCTGAGGACACTCTGCTACGAGCCCTTGTGAG 1620
1b 1621 CATTCGGAAGATTTTCGACTACATTTGATGATTAATTCAGATTTGATGATTCAT 1680
1b 1621 CATTCGGAAGATTTTCGACTACATTTGATGATTAATTCAGATTTGATGATTCAT 1680
1b 1681 GATGACACAAATCCTCATCTGGGACTTCTTAATGATCCAGTGGCCAGCTGAACCCCCC 1740
1b 1681 GATGACACAAATCCTCATCTGGGACTTCTTAATGATCCAGTGGCCAGCTGAACCCCCC 1740
1b 1741 CGTTCCCTTCTCGAACATACACCTACATCTCCAGATAAATACATACATGACCTCAT 1800
1b 1741 CGTTCCCTTCTCGAACATACACCTACATCTCCAGATAAATACATACATGACCTCAT 1800
1b 1801 ACTTGGCCAGGACCCATTAAGTTGGGATTTAAGTATCTGCAATACAGGATGAGC 1860
1b 1801 ACTTGGCCAGGACCCATTAAGTTGGGATTTAAGTATCTGCAATACAGGATGAGC 1860

```

RESULT 5

AA84610
ID AAC84610 standard; DNA; 2151 BP.

XX AAC84610;

XX AC

XX 02-APR-2001 (first entry)

XX Human ZP11 protein encoding DNA.

XX S-phase kinase associated protein; SKP1; SKP2; SKP2-like protein; ZF;

XX CUL-1; cullin; CDC53; p27; cyclin E; Max; Mad; c-Myc; MDM2; p53; Bax;

XX Bad; Bcl-2; tumour; cytosolic; ds.

XX Homo sapiens.

XX WO200075184-A1.

XX 14-DEC-2000;

XX 05-JUN-2000; 2000WO-US15449.

XX 04-JUN-1999; 99US-0137494.

XX (UYVA) UNIV YALE.

XX Zhang H, Tsvetkov LM, Kondo T;

XX WPI; 2001-061703/07.

XX P-PSDB; AAB48298.

XX Modulating polypeptide levels in a cell, diagnosing and treating tumor,

XX involves altering levels of proteins such as S-phase kinase associated

XX proteins 1, 2 and cullin/CDC53 proteins -

XX Examples; Page 129-130; 162pp; English.

XX The invention relates to methods of altering the polypeptide levels in a

XX cell, using proteins selected from S-phase kinase associated proteins 1

XX and 2 (SKP1, SKP2), SKP2-like proteins (ZF) and CUL-1 (a member of the

XX cullin/CDC53 family of proteins). The method is useful for altering the

XX level of p27, cyclin E, Max, Mad, c-Myc, MDM2, p53, Bax, Bad or Bcl-2

XX polypeptide in a cell, SKP2 and SKP2-like protein levels are useful for

XX detecting tumours, and in monitoring tumor treatment in a mammal. Agents

XX that modulate interactions between SKP and target proteins are useful for

XX treating tumours.

XX Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;

XX Query Match 100.0%; Score 2151; DB 22; Length 2151;

XX Best Local Similarity 100.0%; Pred. No. 0;


```

RESULT 6
VAL41041
CD AAL41041 standard; cDNA; 2151 BP.
CX
CX AAL41041;
CX
CX 11-OCT-2002 (first entry)
CX
CX cDNA of Human F-box protein FBP1 SEQ ID No 1.
CX
CX Cytostatic; immunomodulator; cardiant; antiinflammatory; antimicrobial;
CX proliferative; differentiative disorder; Sxp2; F-box protein; cancer;
CX ubiquitin ligase; breast cancer; prostate cancer; ovarian cancer;
CX small cell lung carcinoma; immune disorder; parathyroid adenoma; FBP;
CX inflammatory disorder; lymphoma; major opportunistic infection;
CX certain cardiovascular disease; human; gene; ss.
CX
CX Homo sapiens.
CX
CX WO200255665-A2.
CX
CX 18-JUL-2002.
CX
CX 07-JAN-2002; 2002WO-US00311.
CX
CX 05-JAN-2001; 2001US-260179P.
CX
CX (UUNY ) UNIV NEW YORK STATE.
CX
CX Pagano M;
CX
CX WPI; 2002-59965/64.
CX
CX P-PSDB; AAO22446.
CX
CX Screening compounds for treating proliferative disorders, e.g. breast
CX cancer or prostate cancer, infections or immune disorders, comprises
CX detecting a change in the activity of Sxp2 with either p27 or Cks1 -
CX
CX Disclosure; Fig 3; 246pp; English.
CX
CX The invention relates to screening compounds useful for the treatment of
CX proliferative or differentiative disorders comprising detecting a change
CX in the activity of Sxp2 (F-box protein). The method is useful for
CX screening compounds for the treatment of proliferative or differentiative
CX disorders, particularly cancer. These compounds include small molecules,
CX or compounds or derivatives or analogues of the new ubiquitin ligases.
CX The compounds are useful for treating diseases such as cancer (e.g.
CX breast cancer, prostate cancer or ovarian cancer, lymphoma, small cell
CX lung carcinoma or parathyroid adenomas), major opportunistic infections,
CX immune disorders, certain cardiovascular diseases or inflammatory
CX disorders. This polynucleotide sequence represents the cDNA encoding an
CX F-box protein (FBP) of the invention.
CX
CX Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;
CX
CX Query Match 100.0%; Score 2151; DB 24; Length 2151;
CX Best Local Similarity 100.0%; Pred. No. 0;
CX Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CX
CX 1 TCGGTTGGCTGGCGCTGGCCACCAAGGGGGGGCCCGCGGAGAGCGGACCCAGTGGCC 60
CX
CX 1 TCGGTTGGCTGGCGCTGGCCACCAAGGGGGGGCCCGCGGAGAGCGGACCCAGTGGCC 60
CX
CX 61 TCGGCGATTATGACCCCGCGGCGGCTGTCGAAGAGAGGCGACTCAAGTTATGAAT 120
CX
CX 61 TCGGCGATTATGACCCCGCGGCGGCTGTCGAAGAGAGGCGACTCAAGTTATGAAT 120
CX
CX 121 TCCTCAGAGAGAGAGACTGTATATATGGCGAACCCCTAGGAAGATATACCAAGAGAG 180
CX
CX 121 TCCTCAGAGAGAGAGACTGTATATGGCGAACCCCTAGGAAGATATACCAAGAGAG 180
CX
CX 181 AATTCACTTAGACAGACATACACAGCTGTGCCAGACTCTGTTAAACCAAGAAACAGTA 240

```

```

181 AATTCACTTAGACAGACATACACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTA 240
241 TGTTTAGCAAGCACTGCTATGAAGAGCTGAGAAATTGTGTGCCCAAAACAAACCTTGCAT 300
241 TGTTTAGCAAGCACTGCTATGAAGAGCTGAGAAATTGTGTGCCCAAAACAAACCTTGCAT 300
301 GGCACTTCCAGTATGATGTGCCCAAGCAACGGAACCTCTCAGCAAGCTATGAAAGGAA 360
301 GGCACCTTCCAGTATGATGTGCCCAAGCAACGGAACCTCTCAGCAAGCTATGAAAGGAA 360
361 AAGGAACCTGTGTGCAATATCTTGGAGCACTGTGAGAGCTCAGATCAAGTGAATTTGG 420
361 AAGGAACCTGTGTGCAATATCTTGGAGCACTGTGAGAGCTCAGATCAAGTGAATTTGG 420
421 GAAACATCTTATATCCCAAAATGTGTATACCAACATGGGCAATAAACCTCGTATCTTAA 480
421 GAAACATCTTATATCCCAAAATGTGTATACCAACATGGGCAATAAACCTCGTATCTTAA 480
481 CCTATGTTGAGAGAGATTTTCAATAGCTGTCTGCCAGCTCGGGATTTGGATCATATCGCT 540
481 CCTATGTTGAGAGAGATTTTCAATAGCTGTCTGCCAGCTCGGGATTTGGATCATATCGCT 540
541 GAGAACATTTCTGTATACCTGGATGCCAAATCACTATGTCTGCTGCTGCTGCTGCTGCT 600
541 GAGAACATTTCTGTATACCTGGATGCCAAATCACTATGTCTGCTGCTGCTGCTGCTGCT 600
601 GAATGGTACCGAGTGACCTCTGATGGCATGCTGTGGAGAGAGCTTATCGAGAGATGGTC 660
601 GAATGGTACCGAGTGACCTCTGATGGCATGCTGTGGAGAGAGCTTATCGAGAGATGGTC 660
661 AGGACAGATTTCTGTGGAGAGGCTGGCAGAACGAGAGAGGATGGGACAGTATTTATTC 720
661 AGGACAGATTTCTGTGGAGAGGCTGGCAGAACGAGAGAGGATGGGACAGTATTTATTC 720
721 AAAAACAACCTCTGTGAGGGGATGCTCTCCCAACTCTTTTATAGAGACATTTATTC 780
721 AAAAACAACCTCTGTGAGGGGATGCTCTCCCAACTCTTTTATAGAGACATTTATTC 780
781 AAAATTATACAAGACATTTGAGACATAGATCTAATTGGAGATGTGGAGACATAGTTTA 840
781 AAAATTATACAAGACATTTGAGACATAGATCTAATTGGAGATGTGGAGACATAGTTTA 840
841 CAGAGAAATCTACCTCCGAAGTGAACAGCAAGAGAGATTTACTGTTTACAGTATGAT 900
841 CAGAGAAATCTACCTCCGAAGTGAACAGCAAGAGAGATTTACTGTTTACAGTATGAT 900
901 CAGAAATAGTAAGCGGCTTTCGAGACAAACACAAATCAAGATCTGGGATAAACACATTG 960
901 CAGAAATAGTAAGCGGCTTTCGAGACAAACACAAATCAAGATCTGGGATAAACACATTG 960
961 GAATGCAAGCGAATTTCAAGGCCATACAGGTTCAAGTTCAGTCTCTGCTCCAGTATGATG 1020
961 GAATGCAAGCGAATTTCAAGGCCATACAGGTTCAAGTTCAGTCTCTGCTCCAGTATGATG 1020
1021 AGAGTGATCATACAGGATCATCGGATTCACGGTTCAGAGTGTGGATTAATAACAGGT 1080
1021 AGAGTGATCATACAGGATCATCGGATTCACGGTTCAGAGTGTGGATTAATAACAGGT 1080
1081 GAAATGCTAAACAGCTTGTATCCATTTGGAAGAGTCTTGCACCTTGCCTTCAATAAT 1140
1081 GAAATGCTAAACAGCTTGTATCCATTTGGAAGAGTCTTGCACCTTGCCTTCAATAAT 1140
1141 GGCATGATGTGACCTCTCCAAAGATCGTTCCATTTGTTATGGATATGGCTCCCA 1200
1141 GGCATGATGTGACCTCTCCAAAGATCGTTCCATTTGTTATGGATATGGCTCCCA 1200
1201 ACTGACATTTACCTCCGAGGGTCTGGTCCGACACCGAGCTGCTCTCAATGTTAGAC 1260
1201 ACTGACATTTACCTCCGAGGGTCTGGTCCGACACCGAGCTGCTCTCAATGTTAGAC 1260
1261 TTTGATGACAGTACATTTGTTGATCTGGGATAGAGTATTAAGGTATGGACACA 1320

```



```

XX WO200153453-A2.
XX
XX 26-JUL-2004
XX
XX 23-DEC-2000; 2000WO-US34960.
XX
XX 21-JAN-2000; 2000US-0488725.
XX
XX 25-APR-2000; 2000US-052317.
XX
XX 09-JUL-2000; 2000US-0598042.
XX
XX 19-JUL-2000; 2000US-0620312.
XX
XX 03-AUG-2000; 2000US-0653450.
XX
XX 14-SEP-2000; 2000US-0662191.
XX
XX 19-OCT-2000; 2000US-0693036.
XX
XX 30-NOV-2000; 2000US-0250583.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
XX Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX Zhao QA, Zhou P, Drmanac RT;
XX
XX WPI: 2001-488707/53.
XX
XX P-PSDB; AM00960.
XX
XX Novel bone-marrow-expressed polynucleotides and polypeptides, useful
XX for treating e.g. cancer and immune deficiency disorders -
XX
XX Claim 1; Page 428; 648pp; English.
XX
XX The present sequence is one of 251 novel human polynucleotides
XX expressed in the bone marrow. The polynucleotide and the
XX polypeptide encoded by it are useful in the treatment of various
XX immune deficiencies and disorders. The deficiencies and disorders may
XX be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
XX infection, or may result from an autoimmune disorder, a coagulation
XX disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
XX suppression of an inflammatory response or treatment of a nervous
XX system disorder such as Alzheimer's disease. Detection of the presence
XX or increased expression of the polynucleotide or the protein it
XX encodes is useful for the diagnosis and/or prognosis of one
XX or more types of cancer. The polynucleotide and polypeptide can be
XX used as nutritional sources or supplements and in the screening of
XX chemical compounds as potential drugs.
XX
XX Sequence 3220 BP; 878 A; 725 C; 770 G; 847 T; 0 other;
XX
XX
XX Query Match 98.6%; Score 2120.8; DB 22; Length 3220;
XX Best Local Similarity 99.8%; Pred. No. 0;
XX Matches 2136; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
XX
XX 1 TCGCTTGGCTGGCCCTGGCCACCAAGGGGGGGCCCGCGGAGAGCGGACCCAGTGGCC 60
XX 59 TCGCTTGGCTGGCCCTGGCCACCAAGGGGGGGCCCGCGGAGAGCGGACCCAGTGGCC 118
XX
XX 61 TCGGCGATTATGACCCGCGGAGGCGGCTGCGAAGAGAGGCACTCAAGTTTATGAAT 120
XX 119 TCGGCGATTATGACCCGCGGAGGCGGCTGCGAAGAGAGGCACTCAAGTTTATGAAT 178
XX
XX 121 TCTCTAGAGAGAGAGACTGTATATATATGCGAACCCTCTAGGAAGATATATACAGAGAAG 180
XX 179 TCTCTAGAGAGAGAGACTGTATATATATGCGAACCCTCTAGGAAGATATATACAGAGAAG 238
XX
XX 181 AATTCTATTAGACAGACATACACACTGTGCGACACTCTCTTAAACCAAGAACAGTA 240
XX 239 AATTCTATTAGACAGACATACACACTGTGCGACACTCTCTTAAACCAAGAACAGTA 298
XX
XX 241 TGTTAGCAAGCACTGCTATGAGACTGAGAAATGTTGGCCCAAAACAAAACCTTGGCAAT 300
XX 299 TGTTAGCAAGCACTGCTATGAGACTGAGAAATGTTGGCCCAAAACAAAACCTTGGCAAT 358
XX
XX 301 GGCACTTCCAGTATGATTGTGGCCCAAGCAACCGAAACTCTCAGCAGACTATGAAAAGAA 360
XX

```

```

Db 359 GGCACCTTCCAGTATGATGTGTGCCCAAGCAACGGAAACTCTCAGCAAGCTATGAAAAGGAA 418
QY 361 AAGGAACCTGTGTCAAAATACTTTGAGCAGTGTGTGAGCAGTGTGTGAGTCAAGTCAAGTGGAAATTTGTG 420
Db 419 AAGGAACCTGTGTCAAAATACTTTGAGCAGTGTGTGAGCAGTGTGTGAGTCAAGTCAAGTGGAAATTTGTG 478
QY 421 GAAACATCTTATATCCCAATGTGTCTATTACCAACATGGGCAATAACTCGTATCTTAAA 480
Db 479 GAAACATCTTATATCCCAATGTGTCTATTACCAACATGGGCAATAACTCGTATCTTAAA 538
QY 481 CCTATGTTGAGAGAGATTTTCAATACTGTCTGCCAGCTCGGGGATTTGGATCATATCGCT 540
Db 539 CCTATGTTGAGAGAGATTTTCAATACTGTCTGCCAGCTCGGGGATTTGGATCATATCGCT 598
QY 541 GAGAACATCTGTCTATACCTGGATGCAATCACTATGTGTCTGCTGAACCTGTGTGCAAG 600
Db 599 GAGAACATCTGTCTATACCTGGATGCAATCACTATGTGTCTGCTGAACCTGTGTGCAAG 658
QY 601 CAATGGTACCGAGTGACCTCTGATGGCATGCTGTGGAAGAACTTATCGAGAGATTTGTC 660
Db 659 CAATGGTACCGAGTGACCTCTGATGGCATGCTGTGGAAGAACTTATCGAGAGATTTGTC 718
QY 661 AGGACAGATTTCTGTGTGAGAGGCTGTGCGAGAAACGAGAGAGATGGGACAGTATTTATTC 720
Db 719 AGGACAGATTTCTGTGTGAGAGGCTGTGCGAGAAACGAGAGAGATGGGACAGTATTTATTC 778
QY 721 AAAAACAAAACCTCTGACGGGAATGCTCTCCCAACTCTTTTATAGAGCACTTTATCCT 780
Db 779 AAAAACAAAACCTCTGACGGGAATGCTCTCCCAACTCTTTTATAGAGCACTTTATCCT 838
QY 781 AAAATTTATACAGACATTTAGACAATAGATCTTAATTTGGAGATTTGGAGACATAGTTTA 840
Db 839 AAAATTTATACAGACATTTAGACAATAGATCTTAATTTGGAGATTTGGAGACATAGTTTA 898
QY 841 CAGAGAAATCACTGCCGAAGTAAACGAAAGAGGATTTACTGTTTACAGTATCATGAT 900
Db 899 CAGAGAAATCACTGCCGAAGTAAACGAAAGAGGATTTACTGTTTACAGTATCATGAT 958
QY 901 CAGAAATATGATAGCGGCTTTCGAGACAAACAAATCAAGATCTGGGATTAACAAACATTTG 960
Db 959 CAGAAATATGATAGCGGCTTTCGAGACAAACAAATCAAGATCTGGGATTAACAAACATTTG 1018
QY 961 GAATGCAAGCGAATTTCTCAGAGGCCATACAGGTTCAAGTCTCTGCTCCAGTATCATGAG 1020
Db 1019 GAATGCAAGCGAATTTCTCAGAGGCCATACAGGTTCAAGTCTCTGCTCCAGTATCATGAG 1078
QY 1021 AGAGTATCATACAGGATCATCGAATTTCCAGGTCAGAGTGTGGGATGTAATACAGGT 1080
Db 1079 AGAGTATCATACAGGATCATCGAATTTCCAGGTCAGAGTGTGGGATGTAATACAGGT 1138
QY 1081 GAAATGCTAAACACGTTGATTTCACCAATTTGGAAGCAGTTTCTGCACTTTCGCTTCAATTAAT 1140
Db 1139 GAAATGCTAAACACGTTGATTTCACCAATTTGGAAGCAGTTTCTGCACTTTCGCTTCAATTAAT 1198
QY 1141 GGCATGATGTGTGACCTGCTCCAAAGATCGTTCCATTTGCTGTATGGGATATGGCCTCCCA 1200
Db 1199 GGCATGATGTGTGACCTGCTCCAAAGATCGTTCCATTTGCTGTATGGGATATGGCCTCCCA 1258
QY 1201 ACTGACATTTACCTCTCGGAGGGTGTGCTGGACACCCGAGCTGCTCTCAATTTGTAGAC 1260
Db 1259 ACTGACATTTACCTCTCGGAGGGTGTGCTGGACACCCGAGCTGCTCTCAATTTGTAGAC 1318
QY 1261 TTTCATGCAAGTACATTTGTTCTGCACTCTGGGATAGNACTATTAAGGTATGGACACA 1320
Db 1319 TTTCATGCAAGTACATTTGTTCTGCACTCTGGGATAGNACTATTAAGGTATGGACACA 1378
QY 1321 AGTACTTGTGAATTTGTAGGACCTTTAAATGGACACAACACGAGGCAATTCCTGTTTGAG 1380
Db 1379 AGTACTTGTGAATTTGTAGGACCTTTAAATGGACACAACACGAGGCAATTCCTGTTTGAG 1438
QY 1381 TACAGGACAGGCTGTGTAGTGTGCTCATCTGACAAACACTATCAGATTTATGGACATA 1440
Db 1439 TACAGGACAGGCTGTGTAGTGTGCTCATCTGACAAACACTATCAGATTTATGGACATA 1498

```

1441	GAATGTGTGCATGTTTACGAGTGTGTAGAAGGCCATGAGGAATTTGGTCGCTGTATTTCGA	1500
1499	GAATGTGTGCATGTTTACGAGTGTGTAGAAGGCCATGAGGAATTTGGTCGCTGTATTTCGA	1558
1501	TTTGATAACAAGAGGATAGTCAGTGGGGCCCTATGATGGAAAAAATTAAAGTCGTGGGATCTT	1560
1559	TTTGATACAAAGAGGATAGTCAGTGGGGCCCTATGATGGAAAAAATTAAAGTCGTGGGATCTT	1618
1561	GTGGCTGCTTTTGAGACCGCGTGCTCTGAGGGAGACCTCTGTCTACGGACCCCTTGTGGAG	1620
1619	GTGGCTGCTTTTGAGACCGCGTGCTCTGAGGGAGACCTCTGTCTACGGACCCCTTGTGGAG	1678
1621	CATTCCGGAAGAGTTTTTCGACTACAGTTTGATGAATTCAGATTGTCCAGTAGTTCACAT	1680
1679	CATTCCGGAAGAGTTTTTCGACTACAGTTTGATGAATTCAGATTGTCCAGTAGTTCACAT	1738
1681	GATGACACAAATCCTCATCTGGGACTTCTTAATGATCCAGCTGCCCAAGCTGAACCCCC	1740
1739	GATGACACAAATCCTCATCTGGGACTTCTTAATGATCCAGCTGCCCAATCTGAACCCCC	1798
1741	CGTTCGCCCTTCTCGAACATACACTACATCTCCAGATAATAACCATACACTGACCTCAT	1800
1799	CGTTCGCCCTTCTCGAACATACACTACATCTCCAGATAATAACCATACACTGACCTCAT	1858
1801	ACTTGCCCGAGACCCATTAAAG-TTCGGGTATTTAACGTATCTGCCAATACCAAGATGAG	1859
1859	ACTTGCCCGAGACCCATTAAAGTTTCGGGTATTTAACGTATCTGCCAATACCAAGATGAG	1918
1860	CAACAACAGTAAACAATCAAACTACTGCCAGTTTCCCTGGACTAGCCGAGGACAGGGCT	1919
1919	CAACAACAGTAAACAATCAAACTACTGCCAGTTTCCCTGGACTAGCCGAGGACAGGGCT	1978
1920	TTGAGACTCCTGTGGGACACAGTTGGTCTGAGATCGGCCCAGAGCGGTCTACTCAGCAC	1979
1979	TTGAGACTCCTGTGGGACACAGTTGGTCTGAGATCGGCCCAGAGCGGTCTACTCAGCAC	2038
1980	AACTGACTGTCTTAGTGCTGCTATCAGAGATGTCTTCTATCAATGTGGAATGATGGAA	2039
2039	AACTGACTGTCTTAGTGCTGCTATCAGAGATGTCTTCTATCAATGTGGAATGATGGAA	2098
2040	CTTTTAAACCTCCCTCCTCCTCCTCTTTCACCTCTGCACCTAGTTTTTCCCACTGGTT	2099
2099	CTTTTAAACCTCCCTCCTCCTCCTCTTTCACCTCTGCACCTAGTTTTTCCCACTGGTT	2158
2100	CCAGACAAAGTGACTATTAAATATATTAGTGTTTTCCAGAA	2143
2159	CCAGACAAAGTGACTATTAAATATATTAGTGTTTTCCAGAA	2202

RESULT 9

AZ29233

[D AAZ29233]

X

1C AAZ29233;

28-FEB-2000 (Friday)

28-FEB-2000 (first entry)
XX

Human cell signalling protein-12 encoding cDNA.

KW Cell signalling protein-12; CSIGP-12

inflammatory disorder; c

arteriosclerosis; Addison'

Homo sapiens

XXI
 honore sapientis;

FD	16-NOV-1999.
XX	
PF	13-MAY-1999; 99MO-US10567.
XX	
PR	13-MAY-1998; 98US-0085343.
PR	26-AUG-1998; 98US-0098010.
XX	
EA	(INCY-) INCYTE PHARM INC..
XX	
PI	Bandman O, Hillman JL, Lal F, Yue H, Tang YT, Patterson C;
PI	Baughn MR, Yang J;
XX	
DR	WPI: 2000-086432/07.
DR	P-PSDB; AAY44249.
XX	
PT	Human cell signaling proteins useful for, e.g. diagnosing cell
PT	proliferative and inflammatory disorders -
XX	
PS	Claim 9; Page 87-88; 90pp; English.
XX	
CC	The present sequence is a cDNA obtained from Incyte clone 3239149 of
CC	COLAUTOI library. It encodes cell signalling protein-12 (CSIGP-12). It
CC	is expressed in musculo-skeletal, gastrointestinal and nervous
CC	tissues. Fragments of CSIGP encoding nucleic acid can be used as
CC	hybridisation probe for detecting CSIGP related sequences or allelic
CC	variants. Recombinant CSIGP can be produced in host cells by transforming
CC	them with genetically engineered vectors. Agonists or antagonists can be
CC	used in the treatment of cell proliferative and inflammatory disorders
CC	associated with decreased or increased CSIGP expression. CSIGP is used in
CC	the diagnosis, prevention and treatment of cell proliferative disorders
CC	like arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory
CC	disorders like AIDS, Addison's disease, multiple sclerosis, etc.
XX	
SQ	Sequence 2419 BP; 671 A; 531 C; 625 G; 590 T; 2 other;
	Query Match 98.3%; Score 2114.8; DB 21; Length 2419;
	Best Local Similarity 99.7%; Pred. No. 0;
	Matches 2119; Conservative 0; Mismatches 7; Indels 0; Gaps 0
Qy	11 GGCGCCTGCGACAAAGGGGGGCCCGGGGGAGAGCGGACCAGTGGCTCGGGATT 70
Db	11 GGCGCCTGGCACCAAGGGGGGCCCGGGGGAGAGCGGACCAGTGGCTCGGGATT 70
Qy	71 TGGAATCCGGSCGAGCGGTGCTGCAGAAGAGGCACCTCAAGTTTATTAATTCCTCAGAGA 130
Db	71 TGGAATCCGGSCGAGCGGTGCTGCAAGAAGAGGCACCTCAAGTTTATTAATTCCTCAGAGA 130
Qy	131 GAGAAGACTGTAAATAATGGCCAAACCCCCTAGGAAGATATATACCAGAGAGAAGATTCACTTA 190
Db	131 GAGAAGACTGTAAATAATGGCCAAACCCCCTAGGAAGATATATACCAGAGAGAAGATTCACTTA 190
Qy	191 GACAGACATACAACAGCTGTGCCAGACTCTGCTTTAAACCAAGAAAACAGTATGTTTAGCAA 250
Db	191 GACAGACATACAACAGCTGTGCCAGACTCTGCTTTAAACCAAGAAAACAGTATGTTTAGCAA 250
Qy	251 GCACGTCTATGAAGACTGAGAAATTGTGTGCCAACAACAACCTTGCCAATGSCACTTCCA 310
Db	251 GCACGTCTATGAAGACTGAGAAATTGTGTGCCAACAACAACCTTGCCAATGSCACTTCCA 310
Qy	311 GTATGATTGTGCCAACGAAACCTCTCAGCAAGCTATGAAAAAGGAAAAAGGAACCTGT 370
Db	311 GTATGATTGTGCCAACGAAACCTCTCAGCAAGCTATGAAAAAGGAAAAAGGAACCTGT 370
Qy	371 GTGTCAATATCTTTGAGCAGTGGTTCAGAGTCAGATCAAGTGGNAATTTGGGAACATCTTA 430
Db	371 GTGTCAATATCTTTGAGCAGTGGTTCAGAGTCAGATCAAGTGGNAATTTGGGAACATCTTA 430
Qy	431 TATCCCAATATGTCATTATACCAAATCATGGGCACATAAATCGTATCTTAACCTATGTTGC 490
Db	431 TATCCCAATATGTCATTATACCACATGGGCACATAAATCGTATCTTAACCTATGTTGC 490
Qy	491 AGAGAGATTTCAATATGCTCTCGCCAGCTCGGGGATGGATCATATCGCTTGAGAAATTC 550

Db 491 AGAGAGATTTCATAAAGTCTGCTGCGAGCTGGGGATTGGATCATATTGCTGAGAACATTC 550
2y 551 TGTATACCTGGATGCCAAATCACTATGTGTGCTGAACTTGTGCAAGGAATGGTACC 610
Db 551 TGTATACCTGGATGCCAAATCACTATGTGTGCTGAACTTGTGCAAGGAATGGTACC 610
2y 611 GAGTGACCTGTGATGCGATGCTGTGGAGAGCTTATCGAGAGATGGTCAGACAGATT 670
Db 611 GAGTGACCTGTGATGCGATGCTGTGGAGAGCTTATCGAGAGATGGTCAGACAGATT 670
2y 671 CTCTGTGGAGAGCCCTGGCAGAACCGAAGAGAGTGGGACAGTATTATTCAAAAACAAC 730
Db 671 CTCTGTGGAGAGCCCTGGCAGAACCGAAGAGAGTGGGACAGTATTATTCAAAAACAAC 730
2y 731 CTCTGAGCGGATGCTCTCCCACTCTTTTATAGAGCACTTATCCTAAATATTATC 790
Db 731 CTCTGAGCGGATGCTCTCCCACTCTTTTATAGAGCACTTATCCTAAATATTATC 790
2y 791 AAGACATTGAGACAATAGAACTCTAATTGGAGATGTGGAAGACATAGTTTACAGAGAAATC 850
Db 791 AAGACATTGAGACAATAGAACTCTAATTGGAGATGTGGAAGACATAGTTTACAGAGAAATC 850
2y 851 ACTGCCGAGTGAACAAGCAAGAGTTTACTGTTTACAGTATGATGATGATGATGATG 910
Db 851 ACTGCCGAGTGAACAAGCAAGAGTTTACTGTTTACAGTATGATGATGATGATGATG 910
2y 911 TAAGCGGCTTCGAGACAAACAAATCAAGATCTGGGATAAAAACACATTTGGAATGCTAA 970
Db 911 TAAGCGGCTTCGAGACAAACAAATCAAGATCTGGGATAAAAACACATTTGGAATGCTAA 970
2y 971 GAATTTCTCAGAGCCATAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1030
Db 971 GAATTTCTCAGAGCCATAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1030
2y 1031 TAACAGGATCATCGGATTCACCGCTCAGAGTGGGATGTAATACAGGTCAAAATGCTAA 1090
Db 1031 TAACAGGATCATCGGATTCACCGCTCAGAGTGGGATGTAATACAGGTCAAAATGCTAA 1090
2y 1091 ACAGTTGATTCACCAATGTGAGCAGTTCTGCACTTGGTTTCAATTAATGGCATGTTG 1150
Db 1091 ACAGTTGATTCACCAATGTGAGCAGTTCTGCACTTGGTTTCAATTAATGGCATGTTG 1150
2y 1151 TGACCTGCTCCAAAGATCGTTTCCATGCTGTTAGGATATGGCTTCCCACTGACATTA 1210
Db 1151 TGACCTGCTCCAAAGATCGTTTCCATGCTGTTAGGATATGGCTTCCCACTGACATTA 1210
2y 1211 CCCTCCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1270
Db 1211 CCCTCCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1270
2y 1271 AGTACATTTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1330
Db 1271 AGTACATTTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1330
2y 1331 AATTGTAAGGACCTTAATAGGACAAACAGGATTCCTGTTGCTGCTGCTGCTGCTGCTGCT 1390
Db 1331 AATTGTAAGGACCTTAATAGGACAAACAGGATTCCTGTTGCTGCTGCTGCTGCTGCTGCT 1390
2y 1391 GGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1450
Db 1391 GGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1450
2y 1451 CATGTTTACAGTGTAGAGGCTATGAGGATTTGGTGGTGTGCTGCTGCTGCTGCTGCTGCT 1510
Db 1451 CATGTTTACAGTGTAGAGGCTATGAGGATTTGGTGGTGTGCTGCTGCTGCTGCTGCTGCT 1510
2y 1511 AGAGATAGTCAAGTGGGCTTATGATGGAATAATTAAGTGGGATTTGCTGCTGCTGCTGCT 1570
Db 1511 AGAGATAGTCAAGTGGGCTTATGATGGAATAATTAAGTGGGATTTGCTGCTGCTGCTGCT 1570
2y 1571 TGGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1630
Db 1571 TGGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1630

QY 1631 GAGTTTTGAGCTACAGTTTGTAGTAATTCAGATTGTCAAGATTGTCAATGATGACACAA 1690
Db 1631 GAGTTTTGAGCTACAGTTTGTAGTAATTCAGATTGTCAAGATTGTCAATGATGACACAA 1690
QY 1691 TCCTCATCTGGGACTTCCCTAAATGATCCAGCTGCCCAAGCTGAACCCCGTTCCCTTT 1750
Db 1691 TCCTCATCTGGGACTTCCCTAAATGATCCAGCTGCCCAAGCTGAACCCCGTTCCCTTT 1750
QY 1751 CTGGAACATACACCTCATCTCCAGATAAATAACCATACATGACCTCATACTTTGCCAG 1810
Db 1751 CTGGAACATACACCTCATCTCCAGATAAATAACCATACATGACCTCATACTTTGCCAG 1810
QY 1811 GACCCATTAAGTTGGGTATTTAAGCTATCTGCCAATACCAGGATGAGCAACACAGTA 1870
Db 1811 GACCCATTAAGTTGGGTATTTAAGCTATCTGCCAATACCAGGATGAGCAACACAGTA 1870
QY 1871 ACAATCAAACTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1930
Db 1871 ACAATCAAACTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1930
QY 1931 GTTGGGACACAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1990
Db 1931 GTTGGGACACAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1990
QY 1991 TCAGTGTCTGCTATCAGAGATGCTTCTATCAATTTGTAATGATGGAATTTAAACCT 2050
Db 1991 TCAGTGTCTGCTATCAGAGATGCTTCTATCAATTTGTAATGATGGAATTTAAACCT 2050
QY 2051 CCCTGCT 2110
Db 2051 CCCTGCT 2110
QY 2111 TGACTTATAAATATATTAGTGTGTTT 2136
Db 2111 TGACTTATAAATATATTAGTGTGTTT 2136

RESULT 10
AAH89966
ID AAH89966 standard; cDNA; 3622 BP.
XX
AC AAH89966;
XX
DT 01-OCT-2001 (first entry)
XX
DE Human bone marrow cDNA, SEQ ID NO: 97.
XX
KW Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
KW immunosuppressive; gene therapy; cytokine cell proliferation;
KW cell differentiation modulator; immune disorder; infection; cancer;
KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss.
XX
OS Homo sapiens.
XX
XX WO200153453-A2.
XX
PD 26-JUL-2001.
XX
XX
XX 23-DEC-2000; 2000WO-US34960.
XX
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0693036.
XX 30-NOV-2000; 2000US-0250583.
XX
XX (HYSE-) HYSEQ INC.

1 Ford JE, Boyle BU, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
 2 Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
 3 Zhao QA, Zhou P, Drmanac RT;
 4 WPI; 2001-488707/53.
 5 P-PSDB; AAM00847.
 6 Novel bone-marrow-expressed polynucleotides and polypeptides, useful
 7 for treating e.g. cancer and immune deficiency disorders -
 8 Claim 1; Page 274-275; 648pp; English.
 9 The present sequence is one of 251 novel human polynucleotides
 10 expressed in the bone marrow. The polynucleotide and the
 11 polypeptide encoded by it are useful in the treatment of various
 12 immune deficiencies and disorders. The deficiencies and disorders may
 13 be genetic, or may be caused by a viral (e.g. HIV), bacterial or fungal
 14 infection, or may result from an autoimmune disorder, a coagulation
 15 disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
 16 suppression of an inflammatory response or treatment of a nervous
 17 system disorder such as Alzheimer's disease. Detection of the presence
 18 or increased expression of the polynucleotide or the protein it
 19 encodes is useful for the diagnosis and/or prognosis of one
 20 or more types of cancer. The polynucleotide and polypeptide can be
 21 used as nutritional sources or supplements and in the screening of
 22 chemical compounds as potential drugs.
 23 Sequence 3622 BP; 961 A; 839 C; 874 G; 948 T; 0 other;

Query Match 93.5%; Score 2011.4; DB 22; Length 3622;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2026; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
 2Y 112 TTTATGATTCCTCAGAGAGAGAACTGTATATGCGGACCCCTAGAGATAATA 171
 2b 261 TTCCAGATTCTTCAGAGAGAGAACTGTATATGCGGACCCCTAGAGATAATA 320
 2Y 172 CAGAGAGAGAAATTCCTTAGACAGACATACAAAGCTGTGCGACACTCTGCTTAACCAA 231
 2b 321 CAGAGAGAGAAATTCCTTAGACAGACATACAAAGCTGTGCGACACTCTGCTTAACCAA 380
 2Y 232 GAAACAGATGTTTAGCAAGCACTCTATGAGACTGAGATTGTGGCCAAACAA 291
 2b 381 GAAACAGATGTTTAGCAAGCACTCTATGAGACTGAGATTGTGGCCAAACAA 440
 2Y 292 CTTGCCAATGGCACTTCCAGTATGATTGCGCCAGCAACGAACTCTCAGCAAGCTAT 351
 2b 441 CTTGCCAATGGCACTTCCAGTATGATTGCGCCAGCAACGAACTCTCAGCAAGCTAT 500
 2Y 352 GAAAGGAAAGAACTGTGTCAAAATCTTTGAGCAGTGGTCCAGACTCAGATCAAGTG 411
 2b 501 GAAAGGAAAGAACTGTGTCAAAATCTTTGAGCAGTGGTCCAGACTCAGATCAAGTG 560
 2Y 412 GAATTGTGGAAACATCTTATATCCAAATGTGTCAATACCAACATGGGCACATAAACTCG 471
 2b 561 GAATTGTGGAAACATCTTATATCCAAATGTGTCAATACCAACATGGGCACATAAACTCG 620
 2Y 472 TATCTTAACCTATGTCAGAGAGATTTCATAACTGCTGCGCAGCTCGGGATTGGAT 531
 2b 621 TATCTTAACCTATGTCAGAGAGATTTCATAACTGCTGCGCAGCTCGGGATTGGAT 680
 2Y 532 CATATCGCTGAGAAACATCTGTCACTACCTGGATGCAAAATCACTATGTGCTGTGAACCTT 591
 2b 681 CATATCGCTGAGAAACATCTGTCACTACCTGGATGCAAAATCACTATGTGCTGTGAACCTT 740
 2Y 592 GTGTGCAAGGAATGTTACCGAGTACCTCTGATGGCATGCTGTGGAAAGACTTATCGAG 651
 2b 741 GTGTGCAAGGAATGTTACCGAGTACCTCTGATGGCATGCTGTGGAAAGACTTATCGAG 800
 2Y 652 AGAATGGTCAGGACAGATTCTCTGTGGAGAGGCTTGGCAGAACGAGAGGATGGGACAG 711
 2b 801 AGAATGGTCAGGACAGATTCTCTGTGGAGAGGCTTGGCAGAACGAGAGGATGGGACAG 860

QY 712 TATTTATTCAAAAACAAACCTCTCTGACGGGAATGCTCTCTCCCAACTCTTTTATAGACA 771
 Db 861 TATTTATTCAAAAACAAACCTCTCTGACGGGAATGCTCTCTCCCAACTCTTTTATAGACA 920
 QY 772 CTTTATCTCTAAAATTTACAAAGACATTTAGACAAATAGAAATCTAAATGGAGATGTCGAAGA 831
 Db 921 CTTTATCTCTAAAATTTACAAAGACATTTAGACAAATAGAAATCTAAATGGAGATGTCGAAGA 980
 QY 832 CATAGTTTACAGAGAAATTCATCCGGAAGTGAAACAAAGCAAGAGGATTTACTGTTTACAG 891
 Db 981 CATAGTTTACAGAGAAATTCATCCGGAAGTGAAACAAAGCAAGAGGATTTACTGTTTACAG 1040
 QY 892 TATGATCATCAGAAATAGTAAAGGCGCTTCGAGACAAACAAATCAGATCTGGGATAA 951
 Db 1041 TATGATCATCAGAAATAGTAAAGGCGCTTCGAGACAAACAAATCAGATCTGGGATAA 1100
 QY 952 AACACATTGGAATGCAAGCGAATTTCTCAGAGGCATACAGGTTTCAGTCTCTCTCCAG 1011
 Db 1101 AACACATTGGAATGCAAGCGAATTTCTCAGAGGCATACAGGTTTCAGTCTCTCTCCAG 1160
 QY 1012 TATGATCAGAGATGATCATACAGGATCATCGATTCACCGTCAGAGTGGGATGTA 1071
 Db 1161 TATGATCAGAGATGATCATACAGGATCATCGATTCACCGTCAGAGTGGGATGTA 1220
 QY 1072 AATACAGGTGAAATGCTTAAACAGGTTGATTTCACCATTTGTAAGCAGTTCTGCACCTGGGT 1131
 Db 1221 AATACAGGTGAAATGCTTAAACAGGTTGATTTCACCATTTGTAAGCAGTTCTGCACCTGGGT 1280
 QY 1132 TTTCAATATGGCATGATGGTGACCTCTCTCCAAAGATCGTTCCATTTGCTGATGGGATAG 1191
 Db 1281 TTTCAATATGGCATGATGGTGACCTCTCTCCAAAGATCGTTCCATTTGCTGATGGGATAG 1340
 QY 1192 GCCTCCCAACTGACATTTACCTCCGAGGGTCTCGTGGACACCGAGCTGCTGTGCAAT 1251
 Db 1341 GCCTCCCAACTGACATTTACCTCCGAGGGTCTCGTGGACACCGAGCTGCTGTGCAAT 1400
 QY 1252 GTTGTAGACTTTGATGACAAAGTACATTTGTTCTGCAATCTGGGGATAGAACTATAAGGTA 1311
 Db 1401 GTTGTAGACTTTGATGACAAAGTACATTTGTTCTGCAATCTGGGGATAGAACTATAAGGTA 1460
 QY 1312 TGGAAACAAAGTACTCTGTGAATTTGTAAGGACCTTAAATGGACACAAACGAGGCAATGGC 1371
 Db 1461 TGGAAACAAAGTACTCTGTGAATTTGTAAGGACCTTAAATGGACACAAACGAGGCAATGGC 1520
 QY 1372 TGTTTGCAGTACAGGACAGGCTGGTAGTGGTGTCTATCTGACAACTATACAGTATA 1431
 Db 1521 TGTTTGCAGTACAGGACAGGCTGGTAGTGGTGTCTATCTGACAACTATACAGTATA 1580
 QY 1432 TGGGACATAGAAATGTTGTGTGATGTTTACGAGTGTGTAAGAGGCGCATGAGCAATTTGGTGGT 1491
 Db 1581 TGGGACATAGAAATGTTGTGTGATGTTTACGAGTGTGTAAGAGGCGCATGAGCAATTTGGTGGT 1640
 QY 1492 TGTTATTCGATTTGATACAAAGAGGATAGTCACTGGGGCTTATGATGGAAAAATTAAGTG 1551
 Db 1641 TGTTATTCGATTTGATACAAAGAGGATAGTCACTGGGGCTTATGATGGAAAAATTAAGTG 1700
 QY 1552 TGGGATCTTGTGGTGTGTTTGGACCCCGTGTCTCTGAGGACACACTCTGTCTAGGGAC 1611
 Db 1701 TGGGATCTTGTGGTGTGTTTGGACCCCGTGTCTCTGAGGACACACTCTGTCTAGGGAC 1760
 QY 1612 CTTGTGGAGCATTTCCGGAAGAGTGTTCGACTACAGTGTGATGAATTTCCAGATTTGTCAGT 1671
 Db 1761 CTTGTGGAGCATTTCCGGAAGAGTGTTCGACTACAGTGTGATGAATTTCCAGATTTGTCAGT 1820
 QY 1672 AGTTACATGATGACAACTCTCATCTGGGACTTCTCTAAATGATCCAGCTGCCAGCT 1731
 Db 1821 AGTTACATGATGACAACTCTCATCTGGGACTTCTCTAAATGATCCAGCTGCCAGCT 1880
 QY 1732 GAAACCCCGCTTCCCTCTCTCGAAACATACACCTACATCTCCAGATATAATTAACATACAC 1791
 Db 1881 GAAACCCCGCTTCCCTCTCTCGAAACATACACCTACATCTCCAGATATAATTAACATACAC 1940
 QY 1792 TGACCTCATACTTGGCCGAGGACCCATTAAG-TTGGCGTATTTAACGTTATCTGCCAATAC 1850

Best Local Similarity 99.6%; Pred. No. 0;			
Matches 2024; Conservative 0; Mismatches 8; Indels 1; Gaps 1;			
Qy	112	TTTATGAATTCCTCAGAGAGAGAACTGTAAATATGGCGAACCCCTAGGAAGATAATA	171
Db	298	TTCCAGAAATTCCTCAGAGAGAGAACTGTAAATATGGCGAACCCCTAGGAAGATAATA	357
Qy	172	CCAGAGAAATTCCTCAGAGAGAGAACTGTAAATATGGCGAACCCCTAGGAAGATAATA	231
Db	358	CCAGAGAAATTCCTCAGAGAGAGAACTGTAAATATGGCGAACCCCTAGGAAGATAATA	417
Qy	232	GAACAGATGTGTAGCAAGCACTGTATGAAGAGATGAGAAATGTGGCCCAAAACAAA	291
Db	418	GAACAGATGTGTAGCAAGCACTGTATGAAGAGATGAGAAATGTGGCCCAAAACAAA	477
Qy	292	CTTGCCCAATGGCACTTCAGATGATGTGGCCCAAGCAACGAAACTCTCAGCAAGCTAT	351
Db	478	CTTGCCCAATGGCACTTCAGATGATGTGGCCCAAGCAACGAAACTCTCAGCAAGCTAT	537
Qy	352	GAAGAGAAAGAACTGTGTCAAAATCTTTGAGCAAGTGTGAGAGTCAAGTCAAGTG	411
Db	538	GAAGAGAAAGAACTGTGTCAAAATCTTTGAGCAAGTGTGAGAGTCAAGTCAAGTG	597
Qy	412	GAATTTGTGGAACATCTATATCCCAATGTCTCATTTACCAATGGCCGACATAACTCG	471
Db	598	GAATTTGTGGAACATCTATATCCCAATGTCTCATTTACCAATGGCCGACATAACTCG	657
Qy	472	TATCTTAACTATGTGAGAGAGATTTTCAAACTGTCTGCCAGCTCGGGGATTTGAT	531
Db	658	TATCTTAACTATGTGAGAGAGATTTTCAAACTGTCTGCCAGCTCGGGGATTTGAT	717
Qy	532	CATATCGCTCAGAACATCTGTATACCTGGATGGCAATCATATGTCGTGTAACCTT	591
Db	718	CATATGCTCAGAACATCTGTATACCTGGATGGCAATCATATGTCGTGTAACCTT	777
Qy	592	GTGTGCAAGGAATGGTACCGAGTGTCTGTATGGCATGCTGTGGAAGAACTTATCGAG	651
Db	778	GTGTGCAAGGAATGGTACCGAGTGTCTGTATGGCATGCTGTGGAAGAACTTATCGAG	837
Qy	652	AGATGGTCAGGACAGATTCCTGTGGAGGCGCTGGCAGACGAGAGGATGGGACAG	711
Db	838	AGATGGTCAGGACAGATTCCTGTGGAGGCGCTGGCAGACGAGAGGATGGGACAG	897
Qy	712	TATTTATCAAAACAAACCTCTGTGCGGGAATGTCTCTCCAACTCTTTTATAGAGCA	771
Db	898	TATTTATCAAAACAAACCTCTGTGCGGGAATGTCTCTCCAACTCTTTTATAGAGCA	957
Qy	772	TATTTATCAAAACAAACCTCTGTGCGGGAATGTCTCTCCAACTCTTTTATAGAGCA	831
Db	958	CTTTATCTTAAATATACAGACATTTGAGACATAGAAATCTAAATGGAGATGGGAAGA	1017
Qy	832	CATAGTTTACAGAGAAATTCATCTCCGAAAGTGAACAAGCAAGAGAGTTTACTGTTTACAG	891
Db	1018	CATAGTTTACAGAGAAATTCATCTCCGAAAGTGAACAAGCAAGAGAGTTTACTGTTTACAG	1077
Qy	892	TATCATCATCAGAAATAGTAAGCGGCTTCGAGACCAACACATCAAGATCTGGATATAA	951
Db	1078	TATCATCATCAGAAATAGTAAGCGGCTTCGAGACCAACACATCAAGATCTGGATATAA	1137
Qy	952	AACACATTGGAATGCAAGCGAATTTCTCAGAGGCCATACAGGTTTCTGTCTCCAG	1011
Db	1138	AACACATTGGAATGCAAGCGAATTTCTCAGAGGCCATACAGGTTTCTGTCTCCAG	1197
Qy	1012	TATCATCAGAGATGATCATTAACAGGATCATCGGATTCACCGTTCAGAGTGGGATGTA	1071
Db	1198	TATCATCAGAGATGATCATTAACAGGATCATCGGATTCACCGTTCAGAGTGGGATGTA	1257
Qy	1172	AATACAGGTGAATGTAAACAGGTTGATTCACCATTTGTAAGCAGGTTCTGCACCTTGGT	1131
Db	1258	AATACAGGTGAATGTAAACAGGTTGATTCACCATTTGTAAGCAGGTTCTGCACCTTGGT	1317
Qy	1132	TTCAATTAATGGCATGTGGTGTGACCTGCTCCAAAGATCGTTCATGCTGTATGGGATATG	1191

Db	1318	TTCAATTAATGGCATGTGTGTGACCTCTCCAAAGATCGTTCATCTGCTGTATGGGATATG	1377
Qy	1192	GCCTCCCAACTGACATTTACCTCCGAGAGGCTGTGGTCGGACACCGAGCTGTCTCAAT	1251
Db	1378	GCCTCCCAACTGACATTTACCTCCGAGAGGCTGTGGTCGGACACCGAGCTGTCTCAAT	1437
Qy	1252	GTTGTAGACTTTGATGACAAGTACATTTGTTCTGCACTCTGGGGATAGAACTATTAAGGTA	1311
Db	1438	GTTGTAGACTTTGATGACAAGTACATTTGTTCTGCACTCTGGGGATAGAACTATTAAGGTA	1497
Qy	1312	TGGAACCAAAAGTACTTGTGAAATTTGTAAGGACCTTAAATGGAACAAACGAGGCAATGCC	1371
Db	1498	TGGAACCAAAAGTACTTGTGAAATTTGTAAGGACCTTAAATGGAACAAACGAGGCAATGCC	1557
Qy	1372	TGTTTGCAGTACAGGACAGGCTGTGTAGTGTGCTCATCTGCAACACACTATCAGATTA	1431
Db	1558	TGTTTGCAGTACAGGACAGGCTGTGTAGTGTGCTCATCTGCAACACACTATCAGATTA	1617
Qy	1432	TGGAACATAGAAATGTGTGCAATTTTACGAGTGTAGAGGCCATGAGGAAATTTGGTGGT	1491
Db	1618	TGGAACATAGAAATGTGTGCAATTTTACGAGTGTAGAGGCCATGAGGAAATTTGGTGGT	1677
Qy	1492	TGTTTGCAGTACAGGACAGGCTGTGTAGTGTGCTCATCTGCAACACACTATCAGATTA	1551
Db	1678	TGTTTGCAGTACAGGACAGGCTGTGTAGTGTGCTCATCTGCAACACACTATCAGATTA	1737
Qy	1552	TGGAATCTTGTGCTGTGTGCAATTTTACGAGTGTAGAGGCCATGAGGAAATTTGGTGGT	1611
Db	1738	TGGAATCTTGTGCTGTGTGCAATTTTACGAGTGTAGAGGCCATGAGGAAATTTGGTGGT	1797
Qy	1612	CTTCTGAGGACATTCGGAAGAGTGTTCGACATACAGTGTGTAAGTTCAGATTTCTCAGT	1671
Db	1798	CTTCTGAGGACATTCGGAAGAGTGTTCGACATACAGTGTGTAAGTTCAGATTTCTCAGT	1857
Qy	1672	AGTTTACATGATGACAAATCTCTGAGTGTGCAATCTCTAAATGATCCAGTGTCCCAAGCT	1731
Db	1858	AGTTTACATGATGACAAATCTCTGAGTGTGCAATCTCTAAATGATCCAGTGTCCCAAGCT	1917
Qy	1732	GAACCCCGGTTTCCCTTCTCGAACATACACATCTCCAGATCTCCAGATAAATACCATACAC	1791
Db	1918	GAACCCCGGTTTCCCTTCTCGAACATACACATCTCCAGATCTCCAGATAAATACCATACAC	1977
Qy	1792	TGACCTCATATCTGCCCAGGACCCATTAAGTGTGGGTATTTAAACGATCTGCCAATAC	1850
Db	1978	TGACCTCATATCTGCCCAGGACCCATTAAGTGTGGGTATTTAAACGATCTGCCAATAC	2037
Qy	1851	CAGGATGAGCAACAAACAGTAAACATCAAACTACTGCCAGTTTCCCTGAGCTAGCCGAGG	1910
Db	2038	CAGGATGAGCAACAAACAGTAAACATCAAACTACTGCCAGTTTCCCTGAGCTAGCCGAGG	2097
Qy	1911	AGCAGGCTTTGAGACTCTCTGTTGGACACAGTTGTGTGCACTCGGCCAGGACGGTCT	1970
Db	2098	AGCAGGCTTTGAGACTCTCTGTTGGACACAGTTGTGTGCACTCGGCCAGGACGGTCT	2157
Qy	1971	ACTCAGACAACTGACGTCTTCACTGCTGATCAAGAGATCTCTTCTATCAATTTGTGAA	2030
Db	2158	ACTCAGACAACTGACGTCTTCACTGCTGATCAAGAGATCTCTTCTATCTTTTGTGAA	2217
Qy	2031	TGATTTGGAATTTTAAACCTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2090
Db	2218	TGATTTGGAATTTTAAACCTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2277
Qy	2091	CCATTTGTTCCAGACAAAGGTGACTTATAAATATATTTTAGTGTGTTTGGCAGAA	2143
Db	2278	CCATTTGTTCCAGACAAAGGTGACTTATAAATATATTTTAGTGTGTTTGGCAGAA	2330

RESULT 13
AAK52701
ID AAK52701 standard; cDNA; 3003 BP.
XX
AC AAK52701;
XX

JI 06-MQV-2001 (first entry)
 KX Human polynucleotide SEQ ID NO 2230.
 KE
 CX
 CW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 CW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 CW tissue growth factor; immunomodulatory; cancer; leukaemia;
 CW nervous system disorder; arthritis; inflammation; ss.
 CX
 CS Homo sapiens.
 CX WO200157190-A2.
 CN
 CD 09-AUG-2001.
 CX
 CF 05-FEB-2001; 2001WO-US04098.
 CX
 CR 03-FEB-2000; 2000US-0496914.
 CR 27-APR-2000; 2000US-0560875.
 CR 20-JUN-2000; 2000US-0598075.
 CR 19-JUL-2000; 2000US-0620325.
 CR 01-SEP-2000; 2000US-0654936.
 CR 15-SEP-2000; 2000US-0663561.
 CR 20-OCT-2000; 2000US-0693325.
 CR 30-NOV-2000; 2000US-0728422.
 CX
 CA (HYSE-) HYSEQ INC.
 CX
 CI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 CX
 DR WPI; 2001-476283/51.
 DR P-PSDB; AAM79568.
 CX
 PT Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 CX
 PS Claim 1; Page 4586-4587; 6221pp; English.
 CX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to
 CC cytokines, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CX were missing at the time of publication.
 CX
 SQ Sequence 3003 BP; 807 A; 703 C; 721 G; 772 T; 0 other;
 Query Match 93.4%; Score 2008.2; DB 22; Length 3003;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 2024; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
 IY 112 TTATGAATTCCTCAGAGAGAGAGACTGTAATATGGCGAACCCCTAGGAAGATAATA 171
 IDB 298 TTCAGAAATTCCTCAGAGAGAGAGACTGTAATATGGCGAACCCCTAGGAAGATAATA 357
 IY 172 CCAGAGAGAAATTCATTAGACAGACATACACAGCTGTGCCAGACTCTGTTAAACCAA 231
 IDB 358 CCAGAGAGAAATTCATTAGACAGACATACACAGCTGTGCCAGACTCTGTTAAACCAA 417
 IY 232 GAAACAGTAGTGTAGCAGAGCACTGCTATGAAGACTGAGAAATGTGGCCCAACAAAA 291
 IDB 418 GAAACAGTAGTGTAGCAGAGCACTGCTATGAAGACTGAGAAATGTGGCCCAACAAAA 477

292 CTTGCCAATGSCACTTCCAGTATGATGTGCCCAAGCAACGGAAACTCTCAGCAAGCTAT 351
 478 CTTGCCAATGSCACTTCCAGTATGATGTGCCCAAGCAACGGAAACTCTCAGCAAGCTAT 537
 352 GAAAGGAAAGGAAGCACTGTGTCAATACTTTGACAGTGTGTGACAGTCAAGTCAAGTG 411
 538 GAAAGGAAAGGAAGCACTGTGTCAATACTTTGACAGTGTGTGACAGTCAAGTCAAGTG 597
 412 GAATTTGTGGAACATCTTATATATCCAAATGTGTCAATTTACCAACATGGGCACATAACTCG 471
 598 GAATTTGTGGAACATCTTATATCCAAATGTGTCAATTTACCAACATGGGCACATAACTCG 657
 472 TATCTTAAACCTATGTTGCAGAGAGATTTTCAATCTCTCCAGCTCGGGATTTGGAT 531
 658 TATCTTAAACCTATGTTGCAGAGAGATTTTCAATCTCTCCAGCTCGGGATTTGGAT 717
 532 CATATCGCTGAGAACATTTCTGTCTATACCTGGATGCCAAATCACTATGTCTGCTGAATTT 591
 718 CATATTTCTGAGAACATTTCTGTCTATACCTGGATGCCAAATCACTATGTCTGCTGAATTT 777
 592 GTGTGCAAGGATGTTACCGAGTGACCTCTGATGGCATGCTGTGGAAGAGCTTATCGAG 651
 778 GTGTGCAAGGATGTTACCGAGTGACCTCTGATGGCATGCTGTGGAAGAGCTTATCGAG 837
 652 AGAATGTGTGAGGACAGATTTCTGTGTGAGAGGCTGTGAGAGAGAGATGGGACAG 711
 838 AGAATGTGTGAGGACAGATTTCTGTGTGAGAGGCTGTGAGAGAGAGATGGGACAG 897
 712 TATTTATTTCAAAACCAACCTCTCTGACGGGAATGCTCTCCCAACTCTTTTATAGAGCA 771
 898 TATTTATTTCAAAACCAACCTCTCTGACGGGAATGCTCTCCCAACTCTTTTATAGAGCA 957
 772 CTTTATCTCTAAATTTATACAAGACATTTGACCAATAGAATCTTAATTTGGAGATGTGGAAGA 831
 958 CTTTATCTCTAAATTTATACAAGACATTTGACCAATAGAATCTTAATTTGGAGATGTGGAAGA 1017
 832 CATAGTTTACAGAGAAATTCATCGCGAGTGAACCAAGCAAGAGCTTTACTGTTTACAG 891
 1018 CATAGTTTACAGAGAAATTCATCGCGAGTGAACCAAGCAAGAGCTTTACTGTTTACAG 1077
 892 TATGATGATCAGAAATAGTAGAGCGGCTTTGAGAGCAACACCAATCAAGATCTTGGGATAAA 951
 1078 TATGATGATCAGAAATAGTAGAGCGGCTTTGAGAGCAACACCAATCAAGATCTTGGGATAAA 1137
 952 AACCATTTGGAATCAGAGCAAGTAATTTCTCAGAGGCCATACAGGTTTCACTGTCTGTCCAG 1011
 1138 AACCATTTGGAATCAGAGCAAGTAATTTCTCAGAGGCCATACAGGTTTCACTGTCTGTCCAG 1197
 1012 TATGATGAGAGAGTGATCATAACAGGATCATCGGATTCACCGTTCAGAGTGTGGATGTA 1071
 1198 TATGATGAGAGAGTGATCATAACAGGATCATCGGATTCACCGTTCAGAGTGTGGATGTA 1257
 1072 AATCAGGTGAAATGTGTAACAGCTTGATTCACCATTTGTAAGCAGTTCTGCACTTGGCT 1131
 1258 AATCAGGTGAAATGTGTAACAGCTTGATTCACCATTTGTAAGCAGTTCTGCACTTGGCT 1317
 1132 TTCATATGTCATGATGTTGACCTGCTCCAAAGATGCTTCCATTTCTGTATGGGATATG 1191
 1318 TTCATATGTCATGATGTTGACCTGCTCCAAAGATGCTTCCATTTCTGTATGGGATATG 1377
 1192 GCCTCCCCCAACTGACATTTACCTTCGGAGGGTCTGGTGGGACACCCAGCTGCTGTCAAT 1251
 1378 GCCTCCCCCAACTGACATTTACCTTCGGAGGGTCTGGTGGGACACCCAGCTGCTGTCAAT 1437
 1252 GTTGTAGACTTTTGTAGCAAGTACATTTCTTCTGCACTTGGGGATAGACTATTAAGGTA 1311
 1438 GTTGTAGACTTTTGTAGCAAGTACATTTCTTCTGCACTTGGGGATAGACTATTAAGGTA 1497
 1312 TGGAAACACAGTACTTGTGAATTTGTAAAGGACCTTAAATGGACACAAACGAGGCAATTGCC 1371
 1498 TGGAAACACAGTACTTGTGAATTTGTAAAGGACCTTAAATGGACACAAACGAGGCAATTGCC 1557
 1372 TGTTTGCAAGTACAGGGACAGGCTGGTGTAGTGGGCTCATCTTGACAACTATCAGATTA 1431

QY	172	CCAGAGAAGAAATTCATTAGACAGACATACACAGCTGTGCCAGACTCTGCTTAAACCAA	231
Db	337	CCAGAGAAGAAATTCGCTTAGACAGACATACACAGCTGTGCCAGACTCTGCTTAAACCAA	396
QY	232	GAACAGATATGTTAGCAGACCTCTATGAGACTGGAATTCGTGGCCAAACAAA	291
Db	397	GAACAGATATGTTAGCAGACCTCTATGAGACTGGAATTCGTGGCCAAACAAA	456
QY	292	CTTGCCAAATGGCACTTCCAGTATGATGTCGCAAGCAACGAACTCTCAGCAAGCTAT	351
Db	457	CTTGCCAAATGGCACTTCCAGTATGATGTCGCAAGCAACGAACTCTCAGCAAGCTAT	516
QY	352	GAACAGAAAGAACTGTGTGCAAACTATTGAGCAGTGTGAGAGTCAAGTG	411
Db	517	GAACAGAAAGAACTGTGTGCAAACTATTGAGCAGTGTGAGAGTCAAGTG	576
QY	412	GAATTTGTGGAACATCTTATATCCCAATGTCATATACCAACATGGGCACATAAATCG	471
Db	577	GAATTTGTGGAACATCTTATATCCCAATGTCATATACCAACATGGGCACATAAATCG	636
QY	472	TATCTTAAACCTATGTTGAGAGAGTTCATAAGCTCTGCCAGCTCGGGGATTCGAT	531
Db	637	TATCTTAAACCTATGTTGAGAGAGTTCATAAGCTCTGCCAGCTCGGGGATTCGAT	696
QY	532	CATATCGCTGAGAACATCTGTATACCTGGATGCCAAATCACTATGTGTCTGAACTT	591
Db	697	CATATCGCTGAGAACATCTGTATACCTGGATGCCAAATCACTATGTGTCTGAACTT	756
QY	592	GTGTGAAGAAATGTTACGAGTACCTCTGATGTCATGTCGTGGAAGAGCTTATCGAG	651
Db	757	GTGTGAAGAAATGTTACGAGTACCTCTGATGTCATGTCGTGGAAGAGCTTATCGAG	816
QY	652	AGAATGTCAGGACAGATTCCTGTGTGAGAGCCCTGGCAGAACGAAAGAGTGGGACAG	711
Db	817	AGAATGTCAGGACAGATTCCTGTGTGAGAGCCCTGGCAGAACGAAAGAGTGGGACAG	876
QY	712	TATTTATTCAAACAACTCTGTCAGGAGTGTCTCTCCCACTCTTTTATAGAGCA	771
Db	877	TATTTATTCAAACAACTCTGTCAGGAGTGTCTCTCCCACTCTTTTATAGAGCA	936
QY	772	CTTTATCTTAAATATACAGACATTTAGACAAATAGAAATCTAATGAGAGTGGGAAGA	831
Db	937	CTTTATCTTAAATATACAGACATTTAGACAAATAGAAATCTAATGAGAGTGGGAAGA	996
QY	832	CATAGTTACAGAGAAATTCACCTGCCAGTGAACAGCAAGAGTTCAGTGTTCAG	891
Db	997	CATAGTTACAGAGAAATTCACCTGCCAGTGAACAGCAAGAGTTCAGTGTTCAG	1056
QY	892	TATGATGATCAGAAATAGTAAGCGCCCTTCGAGACAAACAAATCAAGATCTGGATATA	951
Db	1057	TATGATGATCAGAAATAGTAAGCGCCCTTCGAGACAAACAAATCAAGATCTGGATATA	1116
QY	952	AACACATGGAATGCAAGCGAATCTCACAGCCATACAGGTTCACTCTCTCTCCAG	1011
Db	1117	AACACATGGAATGCAAGCGAATCTCACAGCCATACAGGTTCACTCTCTCTCCAG	1176
QY	1012	TATGATGAGAGATGATCAACAGGATCATCGGATTCACGGTTCAGAGTGTGGATGTA	1071
Db	1177	TATGATGAGAGATGATCAACAGGATCATCGGATTCACGGTTCAGAGTGTGGATGTA	1236
QY	1072	AATACAGGTGAAATGCTAAACAGTGTGATTCACCATTTGGAAGAGTTCGACTTCGGT	1131
Db	1237	AATACAGGTGAAATGCTAAACAGTGTGATTCACCATTTGGAAGAGTTCGACTTCGGT	1296
QY	1132	TTCATTAATGGCATGATGTTGACCTCTCCAAAGATCGTTCCCATTTGTTATGGGATATG	1191
Db	1297	TTCATTAATGGCATGATGTTGACCTCTCCAAAGATCGTTCCCATTTGTTATGGGATATG	1356
QY	1192	GCCTCCCAACTGACATTTACCTCCGAGGGTCTGTCGGACCGAGCTGCTGTCAAT	1251
Db	1357	GCCTCCCAACTGACATTTACCTCCGAGGGTCTGTCGGACCGAGCTGCTGTCAAT	1416

QY	1252	GTGTAGACTTTGATGACAAAGTACATTTGTTCTGCACTGTGGGATAGAACTATAAAGTA	1311
Db	1417	GTGTAGACTTTGATGACAAAGTACATTTGTTCTGCACTGTGGGATAGAACTATAAAGTA	1476
QY	1312	TGGAACACAAAGTACTTGTGAATTTGTFAAGGACTTAAATGGACACAAACGAGGCAATTGCC	1371
Db	1477	TGGAACACAAAGTACTTGTGAATTTGTFAAGGACTTAAATGGACACAAACGAGGCAATTGCC	1536
QY	1372	TGTTTGAGTACAGGACAGGCTGTGTAGTGTGCTCATCTGACACACACTATCAGATTA	1431
Db	1537	TGTTTGAGTACAGGACAGGCTGTGTAGTGTGCTCATCTGACACACACTATCAGATTA	1596
QY	1432	TGGGACATAGAAATGTGTGCTGCTTTTACAGTGTGTAGAGGCCCATGAGAAATTTGTCGT	1491
Db	1597	TGGGACATAGAAATGTGTGCTGCTTTTACAGTGTGTAGAGGCCCATGAGAAATTTGTCGT	1656
QY	1492	TGTTTCGATTTGATTAACAGAGGATAGTCAGTGGGGCTATGATGGAAAAATTAAGTG	1551
Db	1657	TGTTTCGATTTGATTAACAGAGGATAGTCAGTGGGGCTATGATGGAAAAATTAAGTG	1716
QY	1552	TGGGATCTTGTGGCTGCTTTTGGACCCCGTCTCTGACGGGACACTCTGCTACGGACC	1611
Db	1717	TGGGATCTTGTGGCTGCTTTTGGACCCCGTCTCTGACGGGACACTCTGCTACGGACC	1776
QY	1612	CTTGTGAGGATTCGGGAGAGTTCGACTACAGTTCAGTTCAGATTCAGATTCAGT	1671
Db	1777	CTTGTGAGGATTCGGGAGAGTTCGACTACAGTTCAGTTCAGATTCAGATTCAGT	1836
QY	1672	AGTTACATGATGACACAACTCTCATCTGGGACTTCCTAAATGATCCAGCTGCCCAAGCT	1731
Db	1837	AGTTACATGATGACACAACTCTCATCTGGGACTTCCTAAATGATCCAGCTGCCCAAGCT	1896
QY	1732	GAACCCCGGTTCCCGTTCGGAACATACCTACATCTCCAGATTAATTAACCATAC	1791
Db	1897	GAACCCCGGTTCCCGTTCGGAACATACCTACATCTCCAGATTAATTAACCATAC	1956
QY	1792	TGACCTCATACTTGGCCAGGACCCATTAAAGTTGGGTATTTAAC	1836
Db	1957	TGACCTCATACTTGGCCAGGATTCGAAATCGAATATGATACATAAC	2001

Search completed: October 22, 2003, 18:40:07
Job time : 567 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

XX nucleic - nucleic search, using sw model

run on: October 22, 2003, 20:49:00 ; Search time 576 seconds
(without alignments)
10014.850 Million cell updates/sec

Title: US-09-601-168B-1

perfect score: 2151
sequence: 1 tgcgttgctgcgcctggc.....gtttgcagaaaaaaa 2151

scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1792395 seqs, 134090451 residues

Total number of hits satisfying chosen parameters: 3584790

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2151	100.0	2151	12	US-09-832-161-17 Sequence 17, Appl
2	2151	100.0	2151	13	US-10-042-417-1 Sequence 1, Appl
3	1710	79.5	1710	14	US-10-038-010-7 Sequence 7, Appl
4	883.4	41.1	4230	12	US-09-832-161-15 Sequence 15, Appl
5	657	30.5	657	14	US-10-023-530-1 Sequence 1, Appl
6	435.4	20.2	951	9	US-09-764-848-15 Sequence 15, Appl
7	435.4	20.2	951	12	US-10-222-020-15 Sequence 15, Appl
8	435.4	20.2	951	14	US-10-116-016-15 Sequence 15, Appl
9	272	12.6	479	9	US-09-864-761-2778 Sequence 15, Appl
10	208.4	9.7	370	9	US-09-864-761-1942 Sequence 15, Appl
11	180.2	8.4	319	10	US-09-960-352-2934 Sequence 2934, Ap
12	167.8	7.8	524	12	US-10-029-386-4953 Sequence 4953, Ap
13	162.6	7.6	250	12	US-10-029-386-18709 Sequence 18709, A
14	157.4	7.3	475	11	US-09-948-995-15756 Sequence 15756, A
15	124.2	5.8	1620	12	US-10-245-618-13 Sequence 13, Appl
16	124.2	5.8	1659	12	US-10-245-618-7 Sequence 7, Appl

17	124.2	5.8	1770	12	US-10-245-618-17 Sequence 17, Appl
18	124.2	5.8	1881	10	US-09-213-888-20 Sequence 20, Appl
19	124.2	5.8	1881	10	US-09-328-877A-20 Sequence 20, Appl
20	124.2	5.8	1884	12	US-10-245-618-5 Sequence 5, Appl
21	124.2	5.8	2001	10	US-09-213-888-26 Sequence 26, Appl
22	124.2	5.8	2001	10	US-09-328-877A-26 Sequence 26, Appl
23	124.2	5.8	2010	10	US-09-213-888-24 Sequence 24, Appl
24	124.2	5.8	2010	10	US-09-328-877A-24 Sequence 24, Appl
25	124.2	5.8	2063	12	US-10-245-618-34 Sequence 34, Appl
26	124.2	5.8	2124	12	US-10-245-618-1 Sequence 1, Appl
27	124.2	5.8	2255	12	US-10-245-618-35 Sequence 35, Appl
28	124.2	5.8	3550	10	US-09-213-888-1 Sequence 1, Appl
29	124.2	5.8	3550	10	US-09-328-877A-1 Sequence 1, Appl
30	124.2	5.8	3571	10	US-09-213-888-2 Sequence 2, Appl
31	124.2	5.8	3571	10	US-09-328-877A-2 Sequence 2, Appl
32	82.2	3.8	2237	10	US-09-994-485-7 Sequence 7, Appl
33	82.2	3.8	2237	10	US-09-832-292-11 Sequence 11, Appl
34	71.8	3.3	2208	12	US-10-032-585-6658 Sequence 6658, Ap
35	62.4	2.9	2091	14	US-10-128-714-2506 Sequence 2506, Ap
36	62.4	2.9	2091	14	US-10-128-714-2506 Sequence 2506, Ap
37	62.4	2.9	2167	14	US-10-128-714-1506 Sequence 1506, Ap
38	62.4	2.9	2167	14	US-10-128-714-1506 Sequence 1506, Ap
39	62.4	2.9	4167	14	US-10-128-714-506 Sequence 506, App
40	62.4	2.9	4167	14	US-10-128-714-506 Sequence 506, App
41	60	2.8	60	12	US-09-908-975-14380 Sequence 14380, A
42	57	2.6	65	12	US-09-908-975-29374 Sequence 29374, A
43	55.2	2.5	2037	10	US-09-801-368-313 Sequence 313, App
44	53.8	2.5	475	11	US-09-918-995-19598 Sequence 19598, A
45	53.8	2.5	10668	12	US-09-814-353-20029 Sequence 20029, A

ALIGNMENTS

RESULT 1

US-09-832-161-17
; Sequence 17, Application US/09832161
; Publication No. US20030166587A1
; GENERAL INFORMATION:
; APPLICANT: Manning, Anthony M.
; APPLICANT: Mercutio, Frank
; APPLICANT: Amit, Sharon
; APPLICANT: Ben-Neriah, Yinson
; APPLICANT: Davis, Matti
; APPLICANT: Hatzubai, Ada
; APPLICANT: Lavon, Iris
; APPLICANT: Yaron, Avraham
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING ACTIVATION OF
; FILE REFERENCE: NF-KB
; FILE REFERENCE: 860098.427
; CURRENT APPLICATION NUMBER: US/09/832.161
; CURRENT FILING DATE: 1991-04-09
; PRIOR APPLICATION NUMBER: 09/219,060
; PRIOR FILING DATE: 1993-12-16
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 2151
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-832-161-17

Query Match	100.0%	Score	2151	DB 12	Length	2151
Best Local Similarity	100.0%	Pred. No.	0			
Matches 2151	Conservative	0	Mismatches	0	Indels	0
Gaps	0					
QY	1	TGCGTTGGCTTGGCGCTTGGACCAAGGGGCGCCCGCGGAGAGCGAGACCGACGTTGGCC	60			
Db	1	TGCGTTGGCTTGGCGCTTGGACCAAGGGGCGCCCGCGGAGAGCGAGACCGACGTTGGCC	60			
QY	61	TGCGCGATTATGACCCCGCGGAGGGGCTCTCAAGAGAGCGACCTCAAGTTATGAAT	120			
Db	61	TGCGCGATTATGACCCCGCGGAGGGGCTCTCAAGAGAGCGACCTCAAGTTATGAAT	120			

QY	121	TCTTCAGAGAGAGAAGACTGTAAATATATGGCGRAACCCCTTAGAAGAGATAAATCCAGAGAAG	180
DB	121	TCTTCAGAGAGAGAAGACTGTAAATATATGGCGRAACCCCTTAGAAGAGATAAATCCAGAGAAG	180
QY	181	AATTTCACCTTAGACAGACATACAACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTA	240
DB	181	AATTTCACCTTAGACAGACATACAACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTA	240
QY	241	TGTTTATGAAGCACTGCTATAGAGCACTGAGAAATGTGTGCCARAAACAACACTTGCCCAAT	300
DB	241	TGTTTATGAAGCACTGCTATAGAGCACTGAGAAATGTGTGCCARAAACAACACTTGCCCAAT	300
QY	301	GGCACTTCAGTATGATTTGTGCCAAGCAACGGAACCTCTCAGCAAGCTATGAAAGGAA	360
DB	301	GGCACTTCAGTATGATTTGTGCCAAGCAACGGAACCTCTCAGCAAGCTATGAAAGGAA	360
QY	361	AAGGAACCTGTGTCAAAATACTTTGACAGCTGCTCAGAGTCAGATCAAGTGGAAATTTGTG	420
DB	361	AAGGAACCTGTGTCAAAATACTTTGACAGCTGCTCAGAGTCAGATCAAGTGGAAATTTGTG	420
QY	421	GAACATCTTNTATCCAAATGTGCATATACCAATGGGCAATAACTCGTATCTTTAAA	480
DB	421	GAACATCTTNTATCCAAATGTGCATATACCAATGGGCAATAACTCGTATCTTTAAA	480
QY	481	CTATGTTGCAGAGAGATTTTCATAACCTGCTCCAGCTCGGGGATGGCATCATCGCT	540
DB	481	CTATGTTGCAGAGAGATTTTCATAACCTGCTCCAGCTCGGGGATGGCATCATCGCT	540
QY	541	GAGACATCTGTCACTACCTGGATGCCAATCACTATGTCTGCTGAACCTGTGTGCAAG	600
DB	541	GAGACATCTGTGTATCTGGATGCCAATCACTATGTCTGCTGAACCTGTGTGCAAG	600
QY	601	GAATGGTACCGAGTGACCTCTGATGGCATCTCTGGAAGAAGCTTTATCGAGAGAATGGTC	660
DB	601	GAATGGTACCGAGTGACCTCTGATGGCATCTCTGGAAGAAGCTTTATCGAGAGAATGGTC	660
QY	661	AGGACAGATTTCTCTGTGAGAGGCTCGGAGAACGAGAGGATGGGACAGTATTATTTC	720
DB	661	AGGACAGATTTCTCTGTGAGAGGCTCGGAGAACGAGAGGATGGGACAGTATTATTTC	720
QY	721	AAAAACAACTCTGTGACGGAAATGCTCCTCCAACTCTTTTATAGAGCACTTTATCCT	780
DB	721	AAAAACAACTCTGTGACGGAAATGCTCCTCCAACTCTTTTATAGAGCACTTTATCCT	780
QY	781	AAAAATTATACAGACATTTGACACAAATAGAACTTAAATTGGAGATGTGGAAGACATAGTTTA	840
DB	781	AAAAATTATACAGACATTTGACACAAATAGAACTTAAATTGGAGATGTGGAAGACATAGTTTA	840
QY	841	CAGAGAATTCATCGCGAAGTGAACAAAGCAAGAGAGTTTACTGTTTACAGTATGATGAT	900
DB	841	CAGAGAATTCATCGCGAAGTGAACAAAGCAAGAGAGTTTACTGTTTACAGTATGATGAT	900
QY	901	CAGAAAATAGTAAGCGGCTTCGAGACAACACAATCAAGATCTGGGATAAAAACACATTG	960
DB	901	CAGAAAATAGTAAGCGGCTTCGAGACAACACAATCAAGATCTGGGATAAAAACACATTG	960
QY	961	GAATGCAAGCGAATTTCTCAGAGGCCATACAGGTTCACTCTGTCTCCATGTATGTATGAG	1020
DB	961	GAATGCAAGCGAATTTCTCAGAGGCCATACAGGTTCACTCTGTCTCCATGTATGTATGAG	1020
QY	1021	AGAGTGATCATTAACAGGATCATCGGATTTCCACGGTCAGAGTGTGGGATGTAAATACAGGT	1080
DB	1021	AGAGTGATCATTAACAGGATCATCGGATTTCCACGGTCAGAGTGTGGGATGTAAATACAGGT	1080
QY	1081	GAATCTTAAACACGCTTGATTCACCATTTGTGAGCGAGTTCTGCATCTCGGTTTCAATAAT	1140
DB	1081	GAATCTTAAACACGCTTGATTCACCATTTGTGAGCGAGTTCTGCATCTCGGTTTCAATAAT	1140
QY	1141	GGCATGATGGTGAACCTGCTCCAAAGATCGTTTCCATTGCTGTATGGGATATGGCTCCCCA	1200
DB	1141	GGCATGATGGTGAACCTGCTCCAAAGATCGTTTCCATTGCTGTATGGGATATGGCTCCCCA	1200

Qy	1201	ACTGACATTACCTTCGGAGSGTGCTGCTCGHACACCGAGCTGCTGTCOAATGTTGTAGAC	1260
Db	1201	ACTGACATTACCTTCGGAGSGTGCTGCTCGHACACCGAGCTGCTGTCOAATGTTGTAGAC	1260
Qy	1261	TTTGATGACAAGTACATTTGTTCTGCATCTGGGGATAGAACTATAAAGGTATGGAAACACA	1320
Db	1261	TTTGATGACAAGTACATTTGTTCTGCATCTGGGGATAGAACTATAAAGGTATGGAAACACA	1320
Qy	1321	AGTACTGTGGAATTTGTAAAGACCTTTAAATGAGACAAACGAGGCAATTCCTGTTTTCGAG	1380
Db	1321	AGTACTGTGGAATTTGTAAAGACCTTTAAATGAGACAAACGAGGCAATTCCTGTTTTCGAG	1380
Qy	1381	TACAGGGAACGGCTGGTAGTGAGTGCGTCATCTGCAACAACACTATCAGATTATGGACATA	1440
Db	1381	TACAGGGAACGGCTGGTAGTGAGTGCGTCATCTGCAACAACACTATCAGATTATGGACATA	1440
Qy	1441	GAATGTGGTGATGTTTACGAGTGTTAGAAAGGCCATGAGGAATTTGGTGCGTTGTTATCGA	1500
Db	1441	GAATGTGGTGATGTTTACGAGTGTTAGAAAGGCCATGAGGAATTTGGTGCGTTGTTATCGA	1500
Qy	1501	TTTGATTAACAAGGATAGTCAGTGGGGCCTATGATGGAAAAATTTAAAGTGTGGGATCTT	1560
Db	1501	TTTGATTAACAAGGATAGTCAGTGGGGCCTATGATGGAAAAATTTAAAGTGTGGGATCTT	1560
Qy	1561	GTGGCTGTCTTGGACCCCGTGCTCTGAGGAGACACTCTGTCTACGGACCCCTTGTGGAG	1620
Db	1561	GTGGCTGTCTTGGACCCCGTGCTCTGAGGAGACACTCTGTCTACGGACCCCTTGTGGAG	1620
Qy	1621	CAATTCGGGAAGAGTTTTCGACTACAGTTTGATGAATTCAGATTGTCCAGTAGTTCACAT	1680
Db	1621	CAATTCGGGAAGAGTTTTCGACTACAGTTTGATGAATTCAGATTGTCCAGTAGTTCACAT	1680
Qy	1681	GATGACAACAATCCTCATCTGGGACTTCTTAAATGATCCAGCTGCCCAAGCTGAACCCCCC	1740
Db	1681	GATGACAACAATCCTCATCTGGGACTTCTTAAATGATCCAGCTGCCCAAGCTGAACCCCCC	1740
Qy	1741	CGTTCCTCTTCGGAACATACACTACATCTCCAGATAAATACCATACACTGACCTCAT	1800
Db	1741	CGTTCCTCTTCGGAACATACACTACATCTCCAGATAAATACCATACACTGACCTCAT	1800
Qy	1801	ACTGCCAGGACCCATTAAAGTTGCGGTATTTAACTATCTGCCAATACCAAGATGAGC	1860
Db	1801	ACTGCCAGGACCCATTAAAGTTGCGGTATTTAACTATCTGCCAATACCAAGATGAGC	1860
Qy	1861	AACAACAGTAACAATCAAACTACTGCCAGTTTCCCTGGACTAGCGGAGGACGAGGGCTT	1920
Db	1861	AACAACAGTAACAATCAAACTACTGCCAGTTTCCCTGGACTAGCGGAGGACGAGGGCTT	1920
Qy	1921	TGAGACTCTCTGTTGGGACACAGTTGGTGTGCGATCGGGCCAGGACGGTCTACTCAGACACA	1980
Db	1921	TGAGACTCTCTGTTGGGACACAGTTGGTGTGCGATCGGGCCAGGACGGTCTACTCAGACACA	1980
Qy	1981	ACTGACTGCTTCAGTGTGCTTATCAGAAGATGTCTTCTATCAATTTGTGAATGATTGGAAC	2040
Db	1981	ACTGACTGCTTCAGTGTGCTTATCAGAAGATGTCTTCTATCAATTTGTGAATGATTGGAAC	2040
Qy	2041	TTTTTAAACCTCCCTCCCTCCCTTTCACCTCTGCACTAGTTTTTCCCATTTGGTTC	2100
Db	2041	TTTTTAAACCTCCCTCCCTCCCTTTCACCTCTGCACTAGTTTTTCCCATTTGGTTC	2100
Qy	2101	CAGACAAAGGTGACTTTAAATATATTAGTGTGTTTCCAGAAAAA	2151
Db	2101	CAGACAAAGGTGACTTTAAATATATTAGTGTGTTTCCAGAAAAA	2151

RESULT 2
US-10-042-417-1
; Sequence 1, Application US/10042417
; Publication No. US20020123042A1
; GENERAL INFORMATION:
; APPLICANT: Pagano, M.
; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
; TITLE OF INVENTION: PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS

```

: FILE REFERENCE: 5914-090-999
: CURRENT APPLICATION NUMBER: US/10/042,417
: CURRENT FILING DATE: 2002-01-07
: PRIOR APPLICATION NUMBER: 60/260,179
: PRIOR FILING DATE: 2001-01-5
: NUMBER OF SEQ ID NOS: 89
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 1
: LENGTH: 2151
: TYPE: DNA
: ORGANISM: Homo sapiens
JS-10-042-417-1

```

Query Match	100.0%;	Score 2151;	DB 13;	Length 2151;
Best Local Similarity	100.0%;	Fred. No. 0;		
Matches 2151;	Conservative	0;	Mismatches	0;
			Indels	Gaps
2Y	1	TGCGTTGGTCGGCGCTGGCACCAAGAGGGGCGGCCCGGGAGAGCGGACCCAGTGGCC	60	
2b	1	TGCGTTGGTCGGCGCTGGCACCAAGAGGGGCGGCCCGGGAGAGCGGACCCAGTGGCC	60	
2Y	61	TGGCGATTATGGACCCCGCGCGGCTGTCAGAGAGGCACTCAAGTTTATGAAT	120	
2b	61	TGGCGATTATGGACCCCGCGCGGCTGTCAGAGAGGCACTCAAGTTTATGAAT	120	
2Y	121	TCTCAGAGAGAGAGACTGTATAATATGGCGAACCCCTAGGAAGATAATACGAGAAG	180	
2b	121	TCTCAGAGAGAGAGACTGTATAATATGGCGAACCCCTAGGAAGATAATACGAGAAG	180	
2Y	181	AATTCACTTAGACACATACAAACAGCTGTGCCAGACTCTGCTTAAACCAAGAACAGTA	240	
2b	181	AATTCACTTAGACACATACAAACAGCTGTGCCAGACTCTGCTTAAACCAAGAACAGTA	240	
2Y	241	TGTTAGCAAGCACTGTCTATGAAGACTGAGAAATGTGTGCCAAAACAAACTTCCCAAT	300	
2b	241	TGTTAGCAAGCACTGTCTATGAAGACTGAGAAATGTGTGCCAAAACAAACTTCCCAAT	300	
2Y	301	GGCACTTCCAGTATGATTTGTGCCCAAGCAACGGAACTCTCAGCAGCTATCAAAAGAA	360	
2b	301	GGCACTTCCAGTATGATTTGTGCCCAAGCAACGGAACTCTCAGCAGCTATCAAAAGAA	360	
2Y	361	AAGGAACCTGTGTCAAAATCTTTGACGAGTGTGCAGATCAGATCAAGTGGAAATTTGTG	420	
2b	361	AAGGAACCTGTGTGTCAAAATCTTTGACGAGTGTGCAGATCAGATCAAGTGGAAATTTGTG	420	
2Y	421	GAACATCTTATATCCCAAATGTGTCATTAACCATGGGCACATAAATCTGTCATCTTAAA	480	
2b	421	GAACATCTTATATCCCAAATGTGTCATTAACCATGGGCACATAAATCTGTCATCTTAAA	480	
2Y	481	CTGTATGTGCAGAGAGATTCATAAATCTGCTGCCAGCTCGGGATGGGATCATATCGCT	540	
2b	481	CTGTATGTGCAGAGAGATTCATAAATCTGCTGCCAGCTCGGGATGGGATCATATCGCT	540	
2Y	541	GAGAACTTCTGTATACCTGGATGCCAAATCACTATGTGCTGCTGAACTTGTGTGCAAG	600	
2b	541	GAGAACTTCTGTATACCTGGATGCCAAATCACTATGTGCTGCTGCTGCTGCAAG	600	
2Y	601	GAATGTCACGAGTACCTCTGTATGGGATGCTGTGGAGAGGCTTATCAGAGAGATGCTC	660	
2b	601	GAATGTCACGAGTACCTCTGTATGGGATGCTGTGGAGAGGCTTATCAGAGAGATGCTC	660	
2Y	661	AGGACAGATTCTCTGTGGAGAGGCTGTGGAGAACGAGAGGATGGGACAGTATTTATTC	720	
2b	661	AGGACAGATTCTCTGTGGAGAGGCTGTGGAGAACGAGAGGATGGGACAGTATTTATTC	720	
2Y	721	AAAAACAAACCTCTCTGACGGAAATGCTCTCCCAATCTTTTTATAGACATTTATCCT	780	
2b	721	AAAAACAAACCTCTCTGACGGAAATGCTCTCTCCCAATCTTTTTATAGACATTTATCCT	780	
2Y	781	AAATTTATCAAGACATTTGAGACAAATAGATCTTAATTTGGAGATGTGGGAAGCATAGTTTA	840	
2b	781	AAATTTATCAAGACATTTGAGACAAATAGATCTTAATTTGGAGATGTGGGAAGCATAGTTTA	840	

[illegible]

RESULT 3

```

US-10-038-010-7
; Sequence 7, Application US/10038010
; Publication No. US20030040089A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, Legrain
; TITLE OF INVENTION: Protein-protein interactions in adipocyte cells
; FILE REFERENCE: B4767A
; CURRENT APPLICATION NUMBER: US/10/038,010
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 60/259,377
; PRIOR FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1..(1710))
; OTHER INFORMATION: Human beta-Trcp1 : F-box containing protein with 7 WD40 repeats;
; OTHER INFORMATION: Part of SCF (Skp1/Cullin/F-box) complex E3 ubiquitin ligase; Impl
; OTHER INFORMATION: located in the degradation of b-catenin and IkBa
; US-10-038-010-7

```

Query Match	79.5%;	Score 1710;	DB 14;	Length 1710;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1710;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	70	ATGGACCCGGCGGCGGTCTCTCAAGAGCAAGGCACTCAAGTTTATGAAATTCCTCAGAG	129
Db	1	ATGGACCCGGCGGCGGTCTCTCAAGAGCAAGGCACTCAAGTTTATGAAATTCCTCAGAG	60
Qy	130	AGAGAAGACTGTAAATATGTCGCGAAACCCCTTAGGAAGATAATACGAGAGAAGAAATTCACCT	189
Db	61	AGAGAAGACTGTAAATATGTCGCGAAACCCCTTAGGAAGATAATACGAGAGAGAAATTCACCT	120
Qy	190	AGACAGACATACAAACAGCTGTGCCAGACTCTGTTTAAACCAAGAAACAGTATGTTTAGCA	249
Db	121	AGACAGACATACAAACAGCTGTGCCAGACTCTGTTTAAACCAAGAAACAGTATGTTTAGCA	180
Qy	250	AGCACTGCTATGAAGACTTGAGAAATGTGTGCCAAAACAAAACATTGCCAATGGCACTTCC	309
Db	181	AGCACTGCTATGAAGACTTGAGAAATGTGTGCCAAAACAAAACATTGCCAATGGCACTTCC	240
Qy	310	AGTATGATTGTCCCAAGCAACGGAAACTCTCAGCAGCTATGAAAGAGAAAGGAACCTG	369
Db	241	AGTATGATTGTCCCAAGCAACGGAAACTCTCAGCAGCTATGAAAGAGAAAGGAACCTG	300
Qy	370	TGTGTCAAAATCTTTGAGCAGTGTGTCAAGTTCAGATCAAGTGGAAATTTGTGGAAACATCTT	429
Db	301	TGTGTCAAAATCTTTGAGCAGTGTGTCAAGTTCAGATCAAGTGGAAATTTGTGGAAACATCTT	360
Qy	430	ATATCCCAAAATGTGTCATTACCAACATGGGCACATAAAGCTCGTATCTTTAAACCTATGTGTG	489

Db	361	ATATCCCAAAATGTGTCAATTACCAACAATGGGCACTAAAACTCGTATCTTAAACCTATCTTG	420
QY	490	CAGAGAGATTTCATAACTGTCTTCGCCAGCTCGGGATTGGATCATATCGCTGAGAACATT	549
Db	421	CAGAGAGATTTCATAACTGTCTTCGCCAGCTCGGGATTGGATCATATCGCTGAGAACATT	480
QY	550	CTGTTCATACCTGGATGGCAAAATCACTATGTGTCTGTGAACCTGTGTGTGCAAGGAATGTGTAC	609
Db	481	CTGTTCATACCTGGATGGCAAAATCACTATGTGTCTGTGAACCTGTGTGTGCAAGGAATGTGTAC	540
QY	610	CGAGTGCCTCTGATGTGCATCTGTGTGAAGAAAGCTTTATCGAGAGAAATGGTCAGGACAGAT	669
Db	541	CGAGTGCCTCTGATGTGCATCTGTGTGAAGAAAGCTTTATCGAGAGAAATGGTCAGGACAGAT	600
QY	670	TCTCTGTGTGGAGCGCTGGGAGAGCAAGAGAGATGGGACAGATTTATTCAAARAAACAA	729
Db	601	TCTCTGTGTGGAGCGCTGGGAGAGCAAGAGAGATGGGACAGATTTATTCAAARAAACAA	660
QY	730	CCTCTGACGGGAATGCTCCTCCCAACTCTTTTTATAGAGCACCTTTATCCTAAAAATTTATA	789
Db	661	CCTCTGACGGGAATGCTCCTCCCAACTCTTTTTATAGAGCACCTTTATCCTAAAAATTTATA	720
QY	790	CAAGACATTGAGACAATPAGAAATCTAATTCGAGATCTGGAAGACATAGTTTACAGAGAATT	849
Db	721	CAAGACATTGAGACAATPAGAAATCTAATTCGAGATCTGGAAGACATAGTTTACAGAGAATT	780
QY	850	CAGTCCGAAAGTGAACAAGCAAGAGGTTTACTGTTTACAGTATGATGATCAGAAAATA	909
Db	781	CAGTCCGAAAGTGAACAAGCAAGAGGTTTACTGTTTACAGTATGATGATCAGAAAATA	840
QY	910	GTAAGCGGCTTCGAGACAACAATCAAGATCTGGATAAAAACACATTCGAATGCAAG	969
Db	841	GTAAGCGGCTTCGAGACAACAATCAAGATCTGGATAAAAACACATTCGAATGCAAG	900
QY	970	CGAATCTCAGGCGCATACAGGTTCAAGTCTCTGTCTCCAGTATGATGAGAGAGTGATC	1029
Db	901	CGAATCTCAGGCGCATACAGGTTCAAGTCTCTGTCTCCAGTATGATGAGAGAGTGATC	960
QY	1030	ATAACAGGATCATCGGATTCACCGGTACAGAGTGTGGGATGAAATACAGGTGAAATGCTA	1089
Db	961	ATAACAGGATCATCGGATTCACCGGTACAGAGTGTGGGATGAAATACAGGTGAAATGCTA	1020
QY	1090	AACAGGTGATTCACCAATGTGAAGCAGTTCTGCACCTCGGTTTCAATAATGGCATGATG	1149
Db	1021	AACAGGTGATTCACCAATGTGAAGCAGTTCTGCACCTCGGTTTCAATAATGGCATGATG	1080
QY	1150	GTGACCTGCTCCAAAGATCGTCCATTTGTGTATGGGATATGGCCTCCCAACTGCAATT	1209
Db	1081	GTGACCTGCTCCAAAGATCGTCCATTTGTGTATGGGATATGGCCTCCCAACTGCAATT	1140
QY	1210	ACCTTCGGAGGGTGTGTGTGACACACGAGCTGTGTCAATGTGTAGACTTTGATGAC	1269
Db	1141	ACCTTCGGAGGGTGTGTGTGACACACGAGCTGTGTCAATGTGTAGACTTTGATGAC	1200
QY	1270	AAGTACATTTGTTCTGCATCTGGGGATGAAACTATAAGGATATGGAACACAAGTACTTGT	1329
Db	1201	AAGTACATTTGTTCTGCATCTGGGGATGAAACTATAAGGATATGGAACACAAGTACTTGT	1260
QY	1330	GAATTTGTAAGGACCTTAAATGTGACACAACGAGGCAATGCTGTTTTCGACTACAGGAC	1389
Db	1261	GAATTTGTAAGGACCTTAAATGTGACACAACGAGGCAATGCTGTTTTCGACTACAGGAC	1320
QY	1390	AGGCTGGTGTGAGTGCTCATCTGACAACTATCAGATTATGGGACATAGAATGTGGT	1449
Db	1321	AGGCTGGTGTGAGTGCTCATCTGACAACTATCAGATTATGGGACATAGAATGTGGT	1380
QY	1450	GCATGTTTACAGGTGTTAGAGGCCATGAGGAATGGTCGTTGTTATTCGATTTGATAAC	1509
Db	1381	GCATGTTTACAGGTGTTAGAGGCCATGAGGAATGGTCGTTGTTATTCGATTTGATAAC	1440
QY	1510	AAGAGGATGATGACGTGGGGCCTATGATGAAAAAAATTAAGATGTGGGATCTTGTGGCTGCT	1569
Db	1441	AAGAGGATGATGACGTGGGGCCTATGATGAAAAAAATTAAGATGTGGGATCTTGTGGCTGCT	1500

2Y 1570 TTGGACCCCGGCTCTGCGAGGACACTCTGCTACGGACCCCTGTGGAGCATTCGGA 1629
Db 1501 TTGGACCCCGGCTCTGCGAGGACACTCTGCTACGGACCCCTGTGGAGCATTCGGA 1560
2Y 1630 AGAGTTTTCGACTACAGTTTGTATGAATTCAGATGTCAGTAGTTCACATGATGACACA 1689
Db 1561 AGAGTTTTCGACTACAGTTTGTATGAATTCAGATGTCAGTAGTTCACATGATGACACA 1620
2Y 1690 ATCCCTCATCTGGAGCTTCTTAATGATCCAGTCCGCAAGCTGACCCCGGTTCCCT 1749
Db 1621 ATCCCTCATCTGGAGCTTCTTAATGATCCAGTCCGCAAGCTGACCCCGGTTCCCT 1680
2Y 1750 TCTGAAACATACACCTACATCTCCAGATAA 1779
Db 1681 TCTGAAACATACACCTACATCTCCAGATAA 1710

RESULT 4
US-09-832-161-15
; Sequence 15, Application US/09832161
; Publication No. US20030166587A1
; GENERAL INFORMATION:
; APPLICANT: Manning, Anthony M.
; APPLICANT: Mercurio, Frank
; APPLICANT: Amit, Sharon
; APPLICANT: Ben-Neriah, Yinon
; APPLICANT: Davis, Matti
; APPLICANT: Hatzubai, Ada
; APPLICANT: Lavon, Iris
; APPLICANT: Yaron, Avraham
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING ACTIVATION OF
; FILE REFERENCE: 860098.427
; CURRENT APPLICATION NUMBER: US/09/832.161
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 09/210,060
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 4230
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-832-161-15

Query Match 41.1%; Score 883.4; DB 12; Length 4230;
Best Local Similarity 74.8%; Pred. No. 5.2e-272;
Matches 1122; Conservative 0; Mismatches 371; Indels 6; Gaps 1;

QY 298 AATGGCACTTCCAGTATGATTGTGCCCAGCAAGCAAGCAACTCTCAGCAAGCTATGAAAG 357
Db 179 AATGGCACTATCTGTGATCTCTCCAGAAAGGCCATCAGAAAGCAACTATCAAAAA 238
QY 358 GAAAGGAACCTGTGTCAAACTATTTGAGCAGTGTGTCAGAGTCAGATCAAGTGGAAATTT 417
Db 239 GAAAGGAACCTGTGTATTAATATTTTACAGTGTCTGAAATCAGATCAAGTGGAAATTT 298
QY 418 GTGGACATCTTATATCCAAATGTCTATTACCAATGCGGCACATAAACTCGATCTT 477
Db 299 GTGGAAACATCTTATTTTACGAATGTCTATTATCAGCATGACATATTAATCTCTTACCTG 358
QY 478 AAACCTATCTTGCAGAGAGATTTCATACTGTCTGCCAGCTCGGGATTGGATCATATC 537
Db 359 AAGCCATGTGTCAGCGGACTTATATACCGCTTACCAGCAAGCCTTAGATCATATA 418
QY 538 GCTGAGAACATCTGTCTACCTGATGCGCAAACTACATATGTGTGTGAACTTGTGTGC 597
Db 419 GAGAAAAACATCTTTTCTGATCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 478
QY 598 AAGGAATGTCAGGAGTACCTCTGATGGCATGCTGTGGAGAGAGCTTATCGAGAGATG 657
Db 479 AAGGAATGTCAGGAGTACCTCTGATGGAGATGCTTTGGAGAGAGCTGATTTGAGCAAGATG 538

QY 658 GTCAGGACAGATTCTCTGTGAGAGGCTGCGAGAACGAAGAGATGGGACAGTATTTA 717
Db 539 GTACCCATCTGATCCCTATGGAAGGACTTTTCAAAAAGAGAGGGTGGATCATCTG 598
QY 718 TTCAAAAAACAAACCTCCCTGAGCGGAAATGCTCCTCCCAACTCTTTTATAGAGACATTTAT 777
Db 599 TTTAAAAACAGACCCACAGATGGC-----CCTCCAAATTCATTTTATAGTCAATATAC 652
QY 778 CCTAAATTTATCAAGACATTTGAGACAAATAGAAATCTAATTTGGAGATGTGGAAGACATAGT 837
Db 653 CCAAGATTTATCCAGGATATAGAGACTATAGAAATCTAATCTAAGGTGTCTACTTTTACAGTAGAT 712
QY 838 TTACAGAGATTTCACTGCCGAAGTGAACAGCAAGCAAGGATTTTACTGTTTACAGTAGAT 897
Db 713 TTGAGAGGATTCAGTCCGCTCTGAAATATGAAAGGTGTCTACTTTTACAGTAGAT 772
QY 898 GATCAAAAAATAGTAAGCGGCTTTGAGACAAACAATCAAGATCTGGGATAAAAAACACA 957
Db 773 GATGAAAAAATATCAGTGGCTACGAGATAATCTATTAAAGATATGGGATAAAACCCAGC 832
QY 958 TTGGAATCGAAGCGAATTTCTCAGAGGCCATACAGGTTTCACTCTCTCTCCAGTATGAT 1017
Db 833 CTGGAATGTTTGAAGGTGTTAAACAGGACACACAGGCTCTCTCTCTCTCTGCGAGTAGAT 892
QY 1018 GAGAGAGTGAATCAACAGAGATCATCGGATTTCCACGGTTCAGAGTGTGGGATGTAATAACA 1077
Db 893 GAGCGTGTCAATTTGATTAACCTGGCTCTTCAAGATTTCTACGGTGAAGTGTGGATGTGAACAGC 952
QY 1078 GGTGAATGCTAAACACAGTTGATTTCAAGATTTGAAAGAGTCTGCACTTGGCTTCAAT 1137
Db 953 GGTGAATGCTTAAACACAGTTGATTTCAAGATTTGAAAGAGTCTGCACTTGGCTTCAAT 1012
QY 1138 AATGGCATGATGCTGACCTGCTCCAAAGATCTGTCATTTGATGGGATATAGGCTCTC 1197
Db 1013 AATGGCATGATGCTGACCTGCTCCAAAGATCTGTCATTTGATGGGATATAGGCTCTC 1072
QY 1198 CCAACTGACATTTACCTCCGAGGGTGTGTCGGAACCCGAGTGTCTGCAATGTGTA 1257
Db 1073 GCGACCGACATCACTTTACGCGGTGTCTGCTGTCGCGGCTGCGGCTGCAATGTAGTA 1132
QY 1258 GACTTTGATGACAAAGTACATTTGTCATCTGGGATAGAACTATAAGGTATGGAAC 1317
Db 1133 GACTTTGACAAAGTACATTTGTCATCTGGGATAGAACTATAAGGTATGGAAC 1192
QY 1318 ACAAGTACTTTGAAATTTGTAAGACCTTAAATGGACACAAAGAGCAATTCCTGTTG 1377
Db 1193 ACGAGCACTGTGAATTTGTTTCTGACTCTCAATGGGACAAAGCGGGCAATTCCTGCTC 1252
QY 1378 CAGTACAGGACAGCGTGTAGTGTGAGTGGCTCATCTGCAACACACTATCAGATTTGGGAC 1437
Db 1253 CAGTACAGGATCGCTGTTGTTAGTGGATCATCAGATTAATACATAGGCTCTGGAT 1312
QY 1438 ATGAAATGCTGTGATGTTTACGAGTGTTAGAAGGCCATAGGAATTTGTCGCTGTTAT 1497
Db 1313 ATGAAATGCTGTGCTGTTTAAAGAGTCTTAGAGGACATGAAGAAATTTGTCGATCATC 1372
QY 1498 CCAATTTGATACAGAGGATAGTCACTGGGGCTTATCATGCAAAATTAAGTGTGGAT 1557
Db 1373 CGGTTTGAACAAAGAGATTTGTCAGTGGGGCTTATCATGGGAAATTAAGTTTGGGAC 1432
QY 1558 CTGTGGCTGCTTTGGACCCCGGCTCTCTGAGGGACACTCTGTCTACGGACCTTGTG 1617
Db 1433 TTGCAAGTCTCTTGACCTCGAGCCCGAGCAGCAGCATTTGTTTGGCAGCATTTGTTG 1492
QY 1618 GAGCATTCGGAAGAGTTTTTCGACTACAGTTTGAATGAATTCAGATTTGTCAGTAGTTCA 1677
Db 1493 GAAATTTCTGACGCTGTGTTTGGGCTCCAGTTGATGATGATGATGATGATGATGATGATGAT 1552
QY 1678 CATGATGACAAATCTCTGAGTCTCTCTAAATGATCCAGCTGCCCAAGCTGAACCC 1737
Db 1553 CATGATGACATATTTTGAATTTGGGATTTCTTAAATGTGCTCTCCAGTGTGCCAGATGAG 1612

QY 1738 CCGCGTCCCTTCTCGAACATACACCTACATCTCCAGATAAATAACCATACACTGACC 1796
 Db 1613 ACCGGTCTCCCTCCGAACATACACTTACATCTCTAGATAACAGTCTGCACTTTCACC 1671

RESULT 5

US-10-023-530-1
 ; Sequence 1, Application US/10023530
 ; Publication No. US20030007956A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LEGRAND, Pierre
 ; APPLICANT: BENAROUS, Richard
 ; APPLICANT: BLOT, Guillaume
 ; APPLICANT: LASSOT, Irina
 ; TITLE OF INVENTION: PROTEINS THAT INTERACT WITH BETA TRCP
 ; FILE REFERENCE: B4717A
 ; CURRENT APPLICATION NUMBER: US/10/023,530
 ; CURRENT FILING DATE: 2002-04-22
 ; PRIOR APPLICATION NUMBER: 60/256,276
 ; PRIOR FILING DATE: 2000-12-18
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 657
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: gene
 ; LOCATION: (1)..(657)
 ; OTHER INFORMATION: Beta TrCP
 US-10-023-530-1

Query Match 30.58; Score 657; DB 14; Length 657;
 Best Local Similarity 100.0%; Pred. No. 7.8e-200; Mismatches 0; Indels 0; Gaps 0;
 Matches 657; Conservative 0;
 QY 70 ATGACCCGCGCGGCGGTGTCGCAAGAGAGAGGCACTCAAGTTTATGAATTCCTCAGAG 129
 Db 1 ATGACCCGCGCGGCGGTGTCGCAAGAGAGAGGCACTCAAGTTTATGAATTCCTCAGAG 60
 QY 130 AGAAGACCTGTATATGCGACACCCCTAGAGATATACCGAGAGAAATTCACCT 189
 Db 61 AGAAGACCTGTATATGCGCAACCCCTAGAGATATACCGAGAGAAATTCACCT 120
 QY 190 AGACAGACATACACAGCTGTGCGAGACTCTGTTTAAACCAAGAAACAGTATGTTAGCA 249
 Db 121 AGACAGACATACACAGCTGTGCGAGACTCTGTTTAAACCAAGAAACAGTATGTTAGCA 180
 QY 250 AGCACTGCTATGAGACTGAGAAATGTTGCGCCAAACAAACCTTGCCTATGGCACTTCC 309
 Db 181 AGCACTGCTATGAGACTGAGAAATGTTGCGCCAAACAAACCTTGCCTATGGCACTTCC 240
 QY 310 AGTATGATTGTGCCAAGCAACGAAACTCTCAGCAAGCTATGAAAAGGAAAGAACTG 369
 Db 241 AGTATGATTGTGCCAAGCAACGAAACTCTCAGCAAGCTATGAAAAGGAAAGAACTG 300
 QY 370 TGTCTCAATCTTGGAGCTGGTGCAGAGTCAGATCAGATCGAATTTGTTGGAACCTT 429
 Db 301 TGTCTCAATCTTGGAGCTGGTGCAGAGTCAGATCAGATCGAATTTGTTGGAACCTT 360
 QY 430 ATATCCCAATGTGCTATTAACCAATGGGCACATAAATCTGATCTTAAACCTATGTTG 489
 Db 361 ATATCCCAATGTGCTATTAACCAATGGGCACATAAATCTGATCTTAAACCTATGTTG 420
 QY 490 CAGAGAGATTTCAATCTGCTGCCAGCTCGGGATTCGATATATCGTCGAGAACTT 549
 Db 421 CAGAGAGATTTCAATCTGCTGCCAGCTCGGGATTCGATATATCGTCGAGAACTT 480
 QY 550 CTGTACACCTGGATGCAATCACTATGTGCTGTGAATCTTGTGCAAGGAATGGTAC 609
 Db 481 CTGTACACCTGGATGCAATCACTATGTGCTGTGAATCTTGTGCAAGGAATGGTAC 540
 QY 610 CGAGTGACCTCTGATGSCATGCTGTGGAAGAACCTTATCGAGAGAAATGGTCAGGACAGAT 669

Db 541 CGAGTGACCTCTGATGSCATGCTGTGGAAGAACCTTATCGAGAGAAATGGTCAGGACAGAT 600
 QY 670 TCTCTGTGGAGAGGCTTGGCAGAACCAAGAGAGATGGGACAGTATTTATTTCAAAAAC 726
 Db 601 TCTCTGTGGAGAGGCTTGGCAGAACCAAGAGAGATGGGACAGTATTTATTTCAAAAAC 657

RESULT 6

US-09-764-848-15
 ; Sequence 15, Application US/09764848
 ; Patent No. US20020077270A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT208
 ; CURRENT APPLICATION NUMBER: US/09/764,848
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 53
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 15
 ; LENGTH: 951
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (883)
 ; OTHER INFORMATION: n equals a.t.g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (913)
 ; OTHER INFORMATION: n equals a.t.g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (931)
 ; OTHER INFORMATION: n equals a.t.g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (935)
 ; OTHER INFORMATION: n equals a.t.g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (941)
 ; OTHER INFORMATION: n equals a.t.g, or c
 US-09-764-848-15

Query Match 20.28; Score 435.4; DB 9; Length 951;
 Best Local Similarity 72.3%; Pred. No. 1.7e-128;
 Matches 597; Conservative 5; Mismatches 216; Indels 8; Gaps 3;
 QY 298 AATGSCACTTCCAGTATGATTGTGCCCAAGCAAGGAAACTCTCAGCAAGCTATCAAAAG 357
 Db 134 AATGGAACATCATCTGTGATCGTCTCCAGNAAGAGCCCATCAGAGGAACCTATCAAAA 193
 QY 358 GAAAAGGAACCTGTGTGTCAAATACCTTTGAGCAGTGTGTCAGATCAAGTGGAAATTT 417
 Db 194 GAAAAGGAACCTGTGTATTAATATTTTGACCAAGTGTCTGAAATCAGATCAAGTGGAAATTT 253
 QY 418 GTGGAACATCTTATATCCCAATGTGTCTTACCAACATGGGCACATAAATCGTATCTT 477
 Db 254 GTGGAACATCTTATTTCAAGATGTGTATTCAGCATGGACATATTAATCTTACCTG 313
 QY 478 AAACCTATGTTGAGAGAGATTTCAATACGTCTGCCAGCTCGGGGATTCAGATCATATC 537
 Db 314 AAGCCCATGTTGACGCGGAGCTTTATACCGCTTTTACCAGAGCAAGGCTTAGATCACATA 373
 QY 538 GCTGAGAACATTTCTGTATACCTCGGATGCCAAATCACTATGTGCTGCTGAACTTGTGTCG 597
 Db 374 CGAGAAACATTTCTTGTACCTGGATGCCAGTCTCTGTGTGACAGAGCTGGTATGCT 433
 QY 598 AAGGAATGTTACCGAGTGAACCTCTGATGGCACTGTGTGGAAGAGCTTATTCAGAGAAATG 657
 Db 434 AAAGAAATGGCAGGAGTGAATCTCAGAAGGAATGCTTTGGAAGAGAGCTGATTGAAACGAATG 493
 QY 658 GTCAGGACAGATTTCTGTGTGAGAGGCTGGCAGAGAGAGGATGGGAGAGTATTATTA 717

494 GTACGCACTGATCCCTCTATGGAAGGACTTTCAGAAAGAGAGGGTGGGATCAGTACCTG 553
718 TTCAAAAAAACCCTCTGACGGGATGCTCTCTCCCAACTCTTTTATAGAGCACTTTAT 777
554 TTATAAAACAGACCCACAG-----ATGGCCCTCCAAATTCATTTATAGTCAATATAC 607
778 CCTAAATATACAGACAGATTGAGACAATAGAACTAATCGGATGTGGAAGACATAGT 837
608 CCAAGATTTATCCAGGATATAGAGACTATAGAACTAATCTGCGGTGTGGACGACAAAC 667
838 TTACAGAGAATTCACCTGCCGAGTGAACAAGCAAGGAGTTTACTGTTCACAGTATGAT 897
668 TTCCAGAGGATTCAGTGGCGCTCTGAAATAGTAAGGTGTCTACTGTTTACAGTACGAT 727
898 GATCAGAAATAGTAAGCGGCTTCGAGACAACACAATCAAGATCTGGGATTAACAAACA 957
728 GATGAAAAAATATCAGTGGCC-TCGAGATAATCTTAAAGATATGGGATAAACCGC 786
958 TTGGAATGAAGGAATCTCACAGCCATACAGTTCAGTCCCTCTGTCTCCAGTATGAT 1017
787 CTGGAATGTTGAAGGTGTTWACAGACACACARGCTCTGK-CTCTGTCTGCAGTATGAT 845
1018 GAGAGAGTATCATACAGAGATCATCGGATTCAGGTTCAGAGTGTGGATGTAAATACA 1077
846 GAKCGGTGATTTKWTACTGGCTCTTTCAGATCTTACCGNAGAGTGTGGATGTGACACCG 905
1078 GGTGAATGCTAAACAGGTGATTCACCATTTGTGAGCAGTCTCTGC 1123
906 GTGAAAGNTTTAAACACATTTGATCCNCCNATTGANGCTGATTTGC 951

RESULT 7
US-10-222-020-15
Sequence 15, Application US/10222020
Publication No. US20030175739A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PTZ08C2
CURRENT APPLICATION NUMBER: US/10/222,020
CURRENT FILING DATE: 2002-08-16
PRIOR APPLICATION NUMBER: 10/116,016
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: 09/764,848
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/225,757
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/226,868
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/216,647
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,267
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/216,880
PRIOR FILING DATE: 2000-07-07

PRIOR APPLICATION NUMBER: 60/225,270
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/251,869
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/235,834
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/234,274
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/234,223
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/228,924
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/224,518
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,369
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/224,519
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,964
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/241,809
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/249,299
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/236,327
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/241,785
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/244,617
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/225,268
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,368
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/251,856
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/251,868
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/229,344
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/234,997
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: 60/229,343
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,345
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,287
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,513
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/231,413
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/229,509
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/236,367
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/237,039
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,038
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/236,370
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/236,802
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,037
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,040
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/240,960
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/239,935
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/239,937


```

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 19492
; LENGTH: 370
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000252.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.75
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: NT HIT: ABQ14596.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST HUMAN HIT: AL040743.1, EVALUE 1.00e-102
; OTHER INFORMATION: SWISSPROT HIT: Q91854, EVALUE 3.00e-46
; US-09-864-761-19492

Query Match 12.6%; Score 272; DB 9; Length 479;
Best Local Similarity 74.2%; Pred. No 3 5e-76;
Matches 359; Conservative 0; Mismatches 120; Indels 5; Gaps 1;

QY 1146 GATGGTACCTGCTCAAGATGCTTCATTCCTGATGGATATGGCTCCCACTGA 1205
DB 479 GATGGTACCTGCTCAAGACCGCTCCATCTGCTGGACATGGCTTCGGACCGA 420
QY 1206 CATTACCTCCGGAGGCTGTGTGCGACACCGAGCTGTCTCAATGTTAGACTTTGA 1265
DB 419 TATCACTTTACGTGGTGTCTGCTGCTCCACTGAGCTGTGTCAAGTAGTAGACTTTGA 360
QY 1266 TGACAAGTACATTTCTTCATCTCGGATAGAACTATAAGGTATGGACACAGTAC 1325
DB 359 CGAAGATACATCGTGTCTGCTGCTGGTACAGAGACCATCAAGTCTGGACATGAGC 300
QY 1326 TTGTGAATTTTGAAGACCTTTAAATGAGACACAAAGAGGAGCATTTGCTGTTGCA 1385
DB 299 CTGTGAATTTGCTGCTACTCTCAATGGGACAAAGCGAG-----TCTGTCTCCAGTACAG 245
QY 1386 GGACAGCTGGTAGTAGTGGCTCATCTGACACACTATCAGATATGGGACATAGATG 1445
DB 244 GGATGGCTGGTGTGTAGTGATCATCATCATCATCATCATCATCATCATCATCATCAT 185
QY 1446 TGGTGCATGTTTACGAGTGTAGAGGCCATGAGGAATGTGTCCTTGTATTCGATTGA 1505
DB 184 TGGCTCTGTTTAAAGTCTCTAGAGGCATGAGGAATGTGTCCTGATCATCCAGTTGA 135
QY 1506 TACACAGATAGTACAGTGGGCTTATCATGGAATTAAGTGTGGGATCTTGTGTC 1565
DB 124 TACACAGATAGTACAGCGGCTTATGATGGAAATTAAGTGTAGGACTTGCAGGC 65
QY 1566 TGCTTTGGACCCCGTCTCTGCGAGGACACTCTGTCTACGGACCCCTTGGAGCATTC 1625
DB 64 CGCTCTGACCTCGAGCCCGACAGACCATGTGTGTGCGACATGTTGGCAATGTTGGACATTC 5
QY 1626 CGGA 1629
DB 4 TGGA 1

RESULT 10
US-09-864-761-19492/c
; Sequence 19492, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng

```



```
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011387.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.1
; OTHER INFORMATION: EST_HUMAN HIT: B1088274.1, EVALUATION 0.00e+00
; OTHER INFORMATION: NT HIT: G16306497, EVALUATION 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q9UXB1, EVALUATION 5.00e-43
US-10-029-386-18709

Query Match          7.6%; Score 162.6; DB 12; Length 250;
Best Local Similarity 78.3%; Pred. No. 3.1e-41;
Matches 195; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

2Y 1059 AGTGTGGGATGTAATACAGGTGAATGCTAAACAGGTGATTCACCAATGTAAGCAGT 1118
Db 250 AGTGTGGGATGTAAGACACGGGTGAAGTTCCTTAACCAATGATCCACCAATGAGGCTGT 191
2Y 1119 TCTGCACITTCGTTTCAATAATGGGATGATGCTGACCTGCTCCAAAGATGTTCCATTGC 1178
Db 190 ATTGCATTACGTTTCAGCAATGGACTGATGTTGACCTGTTCCAGGACCCCTCAATGC 131
2Y 1179 TGTATGGGATGATGCTCCCAACTGACATTAACCTCCGGAGGGTGTGTCGGACACCG 1238
Db 130 TGTGTGGGATGATGCTTCGCGACCGACATCACTTTACGGCGTGTCTGTGGTGGCCACCG 71
2Y 1239 AGCTGTGTCATGTTGTAGACTTTGATGACAGATGATCTTTCTGATCTGCGGGATG 1298
Db 70 GCGTGGCGTCAATGATGATGACTTTGACGACAGTACATCGTGTGCTGCTGTGTGACAG 11
2Y 1299 AACTATAAA 1307
Db 10 GACCATCAA 2

RESULT 14
US-09-918-995-15756
; Publication 15756, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15756
; LENGTH: 475
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(475)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-15756

Query Match          7.3%; Score 157.4; DB 11; Length 475;
Best Local Similarity 92.1%; Pred. No. 2.4e-39;
Matches 164; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

2Y 1966 GGTCTACTGACCACTGCTGCTTCACTGCTGCTATCAGAGATGCTCTTCAAT 2025
Db 55 GGACNACCCAGCACTGCTGCTTCACTGCTGCTATCAGAAATTTCTTTTATCTTT 114
2Y 2026 GTGAATGATGGAACACTTTAAACCTCCCTCTCTCTCTCTCTCTCTCTCTCTCT 2085
```

```
Db 115 GNGAATGATGGAACTTTGGAGGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 174
QY 2086 TTTTCCCATTTGGTTCCAGACAAAGGTGACTTATAAATATATATTAGTGTTCGCCAGAA 2143
Db 175 TTTTCCCATTTGGTTCCAGACAAAGGTGACTTATAAATATATATTAGTGTTCGCCAGAA 232

RESULT 15
US-10-245-618-13
; Sequence 13, Application US/10245618
; Publication No. US20030143582A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven
; APPLICANT: Strohmaier, Helmo
; APPLICANT: Spruck, Charles
; APPLICANT: Sangfelt, Olle
; TITLE OF INVENTION: HCD4 MODULATES CYCLIN E DEGRADATION
; FILE REFERENCE: TSRI 779.2
; CURRENT APPLICATION NUMBER: US/10/245,618
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/404,116
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/322,947
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1620
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-10-245-618-13

Query Match          5.8%; Score 124.2; DB 12; Length 1620;
Best Local Similarity 50.2%; Pred. No. 2.7e-28;
Matches 340; Conservative 0; Mismatches 328; Indels 9; Gaps 1;

QY 883 TGTTCAGATGATGATCAGAAATAGTAAGCGGCTTCGAGACAAACATCAAGATC 942
Db 655 TGCTTACAGTTTGTGGTAAACCGAATAGTTAGTGGTTCCTGATGACAACTTTAAAGTT 714
QY 943 TGGGATAAAACACATTTGGAATGCAAGCGAATTTCTCACAGGCCATACAGGTTCCAGTC 1002
Db 715 TGGTCAGCGTCAACAGCAATGCTGAGAACATTTAGTGGACATACAGGTGGAGTATGG 774
QY 1003 TGTCTCCAGTATGATGAGAGAGTATCATAAAGAGATCTGGATTCACGGTTCAGATG 1062
Db 775 TCATCACAATGAGAGACAAACATCATCATTTAGTGGATCTACAGATCGGACACTCAAAGT 834
QY 1063 TGGGATGTAATACAGGTGAAATGCTAAACACGTTGATTCACCATTTGCAAGCAGTTCTG 1122
Db 835 TGGATGCAGAGCTGGAGATGTATACACCTTATATGGGCATATTCCTCACTGTGCGT 894
QY 1123 CACTTGGCTTTCAATATGCGCATGATGGTCACTGCTGCTCAAAAGATCGTTTCCATGCTGTA 1182
Db 895 TGTATGCATCTTCATGAAAAAAGAGTTGTTAGCGTTTCGAGATGCCACTCTTTAGGGTT 954
QY 1183 TGGGATATGGCTCCCACTCACTGACATTTACCTCCGAGGGTGTGCTCGGACACCGAGCT 1242
Db 955 TGGGATATGAGACAGGCCAGTGT-----ACATGTTTGTGATGGTCACTGTTGA 1005
QY 1243 GCTGTCAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1302
Db 1006 GCAGTCCGCTGTGTTCAATATGATGAGGAGGGTGTGTTAGTGGGCATATGATGATGATG 1065
QY 1303 ATAAGGTGAGACACAGTACTTGTGAATTTGTAGGACCTTAATGAGACACAAACA 1362
Db 1066 GTAAAGGTGGGATCCAGAGACTGAAAACCTGCTACACAGCTGCGAGGGGATCACTAAT 1125
QY 1363 GGCATTTGCTTTTGGAGTACAGGACAGGCTGGTGTAGTGGCTTCATCTGACACACT 1422
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

M nucleic - nucleic search, using sw model

Run on: October 22, 2003, 18:30:59 ; Search time 147 Seconds
(without alignments)
6458.606 Million cell updates/sec

Title: US-09-601-168B-1

Perfect score: 2151

Sequence: 1 tgcgttgctgcggcctggc.....gtttccgagaaaaaaa 2151

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124.4	5.8	126	3	US-09-172-841-2
2	124.4	5.8	126	4	US-08-951-621-2
3	82.2	3.8	2237	4	US-08-914-999-7
4	67.2	3.1	2481	3	US-08-899-578-1
5	50.8	2.4	2085	2	US-08-283-917-8
6	50.8	2.4	2085	2	US-08-961-716-8
7	50.8	2.3	7218	1	US-08-232-463-14
8	47.8	2.2	2152	1	US-08-188-582-17
9	47.8	2.2	2152	1	US-08-646-715-17
10	39.8	1.9	7218	1	US-08-232-463-14
11	38.6	1.8	433	1	US-08-664-596B-1
12	38.6	1.8	433	1	US-08-738-367-1
13	38.6	1.8	733	1	US-08-738-367-6
14	38.6	1.8	3380	2	US-08-156-425-1
15	36.8	1.7	7518	4	US-09-620-312D-1051
16	36.6	1.7	1613	4	US-08-965-048-3
17	36.6	1.7	18597	3	US-09-318-448-11
18	36.6	1.7	18597	4	US-08-962-665-8
19	36.6	1.7	45989	4	US-08-955-048-6
20	36.4	1.7	1844	4	US-08-620-312D-768
21	36.2	1.7	7886	2	US-08-751-189-2
22	36.2	1.7	7886	2	US-08-060-836-2
23	36.2	1.7	7886	3	US-09-184-445-2
24	35.8	1.7	3434	4	US-09-439-313-476
25	35.8	1.7	5829	4	US-08-439-313-473
26	35.4	1.6	7881	2	US-08-751-189-1
27	35.4	1.6	7881	2	US-09-060-836-1

28	35.4	1.6	7881	3	US-09-184-445-1	Sequence 1, Appli
c 29	35.4	1.6	10881	4	US-09-357-206A-9	Sequence 9, Appli
c 30	35.4	1.6	12286	4	US-09-357-206A-1	Sequence 1, Appli
31	34.8	1.6	3465	4	US-08-914-999-5	Sequence 5, Appli
c 32	34.8	1.6	1664976	4	US-08-916-421B-1	Sequence 1, Appli
33	34.6	1.6	1611	4	US-09-302-769-13	Sequence 13, Appli
34	34.6	1.6	1848	1	US-08-313-553-10	Sequence 10, Appli
35	34.6	1.6	1848	3	US-08-767-993-10	Sequence 10, Appli
36	34.6	1.6	2246	3	US-09-032-742-3	Sequence 3, Appli
37	34.6	1.6	2246	3	US-09-032-742-21	Sequence 21, Appli
38	34.6	1.6	2246	3	US-09-032-742-22	Sequence 22, Appli
39	34.6	1.6	2246	3	US-09-032-742-24	Sequence 24, Appli
40	34.6	1.6	2246	3	US-09-032-742-25	Sequence 25, Appli
c 41	34.6	1.6	5558	4	US-08-961-527-103	Sequence 103, App
c 42	34.6	1.6	319608	4	US-09-539-333D-1	Sequence 1, Appli
c 43	34.6	1.6	319608	4	US-09-679-409-1	Sequence 1, Appli
44	34.4	1.6	649	3	US-09-385-982-336	Sequence 336, App
45	34.2	1.6	3747	4	US-09-690-364-17	Sequence 17, Appli

ALIGNMENTS

RESULT 1
US-09-172-841-2
; Sequence 2, Application US/09172841
; Patent No. 6232081
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeffrey W.
; APPLICANT: Elledge, Stephen J.
; TITLE OF INVENTION: F-BOX PROTEINS AND GENES
; FILE REFERENCE: BCM-03510
; CURRENT APPLICATION NUMBER: US/09/172,841
; CURRENT FILING DATE: 1998-10-15
; EARLIER APPLICATION NUMBER: 08/951,621
; EARLIER FILING DATE: 1997-10-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 126
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-172-841-2
Query Match 5.8%; Score 124.4; DB 3; Length 126;
Best Local Similarity 99.2%; Pred. No. 9.4e-31;
Matches 125; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	511	CTGCCAGCTCGGGATTGGATCATATCGCTGAGAACATTTCTGTACATACCTGGATGCCAA	570
Db	1	CTGCCAGCTCGGGATTGGATCATATTTCTGAGAACATTTCTGTACATACCTGGATGCCAA	60
Qy	571	TCATATGTCTGTGAACTTGTGTCAAGGAATGTACCGAGTGACCTCTGTATGGCATG	630
Db	61	TCATATGTCTGTGAACTTGTGTCAAGGAATGTACCGAGTGACCTCTGTATGGCATG	120
Qy	631	CTGTGG 636	
Db	121	CTGTGG 126	

RESULT 2
US-08-951-621-2
; Sequence 2, Application US/08951621
; Patent No. 6573094
; GENERAL INFORMATION:
; APPLICANT: HARPER, JEFFREY W.
; APPLICANT: ELLEDGE, STEPHEN J.
; TITLE OF INVENTION: F-BOX GENES AND PROTEINS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/951,621
FILING DATE: 16-OCT-1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: MacKnight, Kamrin T.
REGISTRATION NUMBER: 38,230
REFERENCE/DOCKET NUMBER: BCM-02999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 126 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"
JS-08-951-621-2

Query Match
Best Local Similarity 5.8%; Score 124.4; DB 4; Length 126;
Matches 125; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2Y 511 CTGCCAGCTCGGGGATGGATCATATCGTCAGAACATTCGTGCATACCTGGATGCCAAA 570
1 CTGCCAGCTCGGGGATGGATCATATTCGTGCAGAACATTCGTGCATACCTGGATGCCAAA 60

2Y 571 TCACATGTGCTGCTGAACCTGTGTCAGGAATGTCACCGAGTGCACCTCTGATGCGATG 630
61 TCACATGTGCTGCTGAACTGTGTCAGGAATGTCACCGAGTGCACCTCTGATGCGATG 120
631 CTGTGG 636
121 CTGTGG 126

RESULT 3
JS-08-914-999-7
Sequence 7, Application US/08914999
Patent No. 6346406
GENERAL INFORMATION:
APPLICANT: Rvazanov, Alexey G.
APPLICANT: Hait, William N.
APPLICANT: Pavuk, Karen S.
TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)
TITLE OF INVENTION: AND METHODS OF USE THEREFOR
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/914,999

FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 601-1-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2237 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Dictyostelium discoideum
US-08-914-999-7

Query Match
Best Local Similarity 3.8%; Score 82.2; DB 4; Length 2237;
Matches 286; Conservative 0; Mismatches 308; Indels 9; Gaps 1;

QY 959 TGGAAATGCAAGCGAATTCACAGCCATACAGGTTCAAGTTCCTCTGTCTCCAGTATGATG 1018
DB 1490 TGGAAATGTTTCAACCTTGAAAGGTTCATGAAGTCCAGTTCGAATCAATTTTATATG 1549

QY 1019 AGAGATGATCATACAGGATCATCGGATTCACCGTCCAGGTGAGTGGGATGTAATACAG 1078
DB 1550 ATCAATATTTGTTTGTGGTTCATCAGATCATCAATTAAGTTTGGGATTTAAAGAAAT 1609

QY 1079 GTGAAATGCTAAACACGCTTGATTCACCATTTGTAAGCAGTTCCTGCACCTTGGCTTTCAATA 1138
DB 1610 TAAGATGATTTTACTTTGGAGGTTCATGATAAAGTTCATACCGTTCCTATTGAATG 1669

QY 1139 ATGCGATGATGTCACCTGCTCCAAAGATCGTTCCATTCGCTGTATGGATATGCGCTCCC 1198
DB 1670 ATAAATATTTGTTTGTGGTTCCTCTGACAAACATATCAAAAGTTTGGGATTTGAAACATTT 1729

QY 1199 CAATGATCATTAACCTCCGAGGGTCTGTCGGACACCGAGTCTGCTCAATGTTGTAG 1258
DB 1730 TGGAAATGTAATATACCTTGAAAGTTCATGCCAGA-----GCCGTCAAACACATTT 1780

QY 1259 ACITTGATGACAAGTACATGTTCTGCACTCTGGGATAGAACTATAAGGTATGGAACA 1318
DB 1781 GTATATCTGTGCAATATTTATTTAGTGGTTCAAATGATAAAACTATCAAGGTTTGGGATTT 1840

QY 1319 CAAGTACTTGTGAATTTGTAAGGACCTTAAATGGACACAAAGAGGSCATTGCTCTTTGTC 1378
DB 1841 TGAACACTTTTCGTTGTAACACTACACTCTAAAGGTTCATACTAAATGGTCCACCACTATCT 1900

QY 1379 AGTACAGGACAGGCTGGTAGTAGTGGCTCTCTGACAACTATCAGATATGGAACA 1438
DB 1901 GTATATTAGTACCAATCTACAGTGGCTCTCTATGATAAAACTATAAGAGTTTGGGATTT 1960

QY 1439 TAGAATGTTGTGATGTTTACGAGTGTAGAAAGGCGCATGAGGAATTTGGTCCGCTGTATTTC 1498
DB 1961 TAAAGAGTTTAGAATGTTCCGCTACTTTAAGAGGCCATGATAGATGGGTGGAACATG 2020

QY 1499 GATTGTGATAACAGAGGATGATGTCAGTGGGCGCTATGATGGAAGAAATTTAAAGTTGGGATC 1558
DB 2021 TAATTTGTGATAAATATTATTACTGTCTAGTACGATAAATAACAATTTAAATTTGGGATTT 2080

QY 1559 TTG 1561
DB 2081 TAG 2083

RESULT 4
US-08-899-578-1
Sequence 1, Application US/08899578

Patent No. 6087153

GENERAL INFORMATION:

APPLICANT: Greenwald, Iva
APPLICANT: Hubbard, E. Jane
TITLE OF INVENTION: SEL-10 AND USES THEREOF
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1195 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,578
FILING DATE: 24-JUL-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/53200/JPW/AKC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 278-0525

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 2481 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

NAME/KEY: CDS

LOCATION: 91...1854

US-08-899-578-1

Query Match 3.1%; Score 67.2; DB 3; Length 2481;

Best Local Similarity 50.6%; Pred. No. 3.9e-11;

Matches 162; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

2Y 1273 TACATTTGTTCTGCACTCTGGGATAGAACTATAAGGTATGGAACACAACTCTGTGAA 1332

Db 1018 TATATTGTTAGCGGTCCACTGATAGAACTGTAAAGTTGGAGTACTGTAGATGGTTCA 1077

2Y 1333 TTGTGAAGGACCTTAATGAGACAAACGAGGCATTGCCTGTTTGCAGTACAGGACAGG 1392

Db 1078 CTCTCTACACTTCAAGGACATACCTCCACTGTTGATGCATGGCTATGGTGTCTC 1137

2Y 1393 CTGTGAGTGAAGTGTCTCTGACAACTATCAGATTATGGGACATAGAAATGTGTGCA 1452

Db 1138 ATACTTGTCACTGGATCAGAGATACCACTCTCGTGTATGGGACGTAGATCCGGACGT 1197

2Y 1453 TGTTACAGAGTTAGAGGCGCATGAGAAATGTGGTGTGTTATTCGATTGTATACAG 1512

Db 1198 CACCTGGCAACTTACATGGCCCATCATGACCCGTTGATCGTTCATTCATTCGATGAACA 1257

2Y 1513 AGGATAGTCAGTGGGGCCCTATGATGGAAAAATTAAGTGTGGATCTTGTGGTGTCTTG 1572

Db 1258 ACTGTGTTTCGGAGAGATATGATTTTACCGTTAAATTTGGAATGCTCATCTGGAGA 1317

2Y 1573 GACCCCGTCTCTGTCAGG 1592

Db 1318 TGTATCCGTACTCTGACCGG 1337

RESULT 5

US-08-283-917-8

, Sequence 8, Application US/08283917

Patent No. 5849557
GENERAL INFORMATION:
APPLICANT: ADACHI, HIDEKI
APPLICANT: TSUJIMOTO, MASAFUMI
APPLICANT: INOUE, KEIZO
APPLICANT: ARAI, HIROYUKI
TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &
ADDRESSEE: NEUSTADT, P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/283,917
FILING DATE: 03-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 209943/1993
FILING DATE: 03-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5849557man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 2292-030-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2085 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Bos taurus
FEATURE:
NAME/KEY: CDS
LOCATION: 844..2073
US-08-283-917-8

Query Match 2.4%; Score 50.8; DB 2; Length 2085;

Best Local Similarity 47.3%; Pred. No. 8.2e-06;

Matches 262; Conservative 0; Mismatches 277; Indels 15; Gaps 3;

QY 1001 TCTCTCCAGTATGATGAGAGATGATCATACAGATCATCGATTCCACGGTCAGAG 1060

Db 1184 TCATTTTCCATCTGTTGTTGATGTTCTGCTTTCAGAGATGCTTACAAATTAAGG 1243

QY 1061 TGTGGGATGTAATACAGGTGAATGCTAAACACGTTGATTCCACATTTGTAAGCAGTTTC 1120

Db 1244 TGTGGGATGATGACACTGGAGATTTTGAACGAACTCTTAAGGGGCATACAGACTCTGTAC 1303

QY 1121 TGCATTCGGTTTCAATTAATGG-----CATGATGGTGACCTGCTCAAGATCGTTCCA 1174

Db 1304 AGGATATTTTTCATTCGACCAAGTGGCAAGCTTCTGGCTTCATGTTCTTCGAGATATGACCA 1363

QY 1175 TTGCTGTATGGGATATGGCTCCCAACTGACATTACCTCCGAGGGTCTCTGTGGAC 1234

Db 1364 TTAAGCTATGGGATTTTCAGGGCTTTGAATGCAATCAGAACCATGCGCATGAC--C 1420

QY 1235 ACCGAGCTGCTGTCAATGTTGTAGACTTTTGATGACAAGTATGTTTTCATCTGGGG 1294

```

;b 1421 ACAATGTTTCTTCAGTAGCCATCATGCCAATGGAGTACATATAGTGTCTGCCTCAAGG 1480
;y 1295 ATAGAACTATAAAGTATGGACACAGTACTTGTGAATTGTGAAGACCTTAAATGGAC 1354
;b 1481 ATAAACATATAAATAATGTGGGAAGTGAACACTGGCTACTGTGTGAAGACATTCACAGGAC 1540
;y 1355 ACAAGAGGAGCATTGCTGTTTGCAGTACAGGGA-----CAGCTGTTAGTGTGCTT 1408
;b 1541 ACAGAGATGGTACGTATGTGTGGCCCAATCAAGACGGCACTTCTATAGCCAGCTGTT 1600
;y 1409 CATCTGACAACTATACAGATATGGAACATAGAAATGTGTGTCATGTTTACAGAGTTAG 1468
;b 1601 CCAATGACAGACTGTGGCTGTATGGTGTGCTGTAGCAACAAGAGGAATGCAAGGCTGAGCTTC 1660
;y 1469 AAGGCCATGAGGAATTTGGTGGCTTCTGATTCGATTTGATACAGAGGATAGTCAGTGGGG 1528
;b 1661 GAGACATGAGCATGTGGTAGNATGCAATTCCTGGGCTCCTGAAAGCTCATATCTTCCA 1720
;y 1529 CCTATGATGGAAAA 1542
;b 1721 TCTCTGAAGCAACA 1734

```

RESULT 6

IS-08-961-716-8

Sequence 8, Application US/08961716

Patent No. 5880272

GENERAL INFORMATION:

APPLICANT: ADACHI, HIDEKI

APPLICANT: TSUJIMOTO, MASAFUMI

APPLICANT: INOUE, KEIZO

APPLICANT: ARAI, HIROYUKI

TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME

NUMBER OF INVENTION: AND GENE THEREOF

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER &

ADDRESSEE: NEUSTADT, P.C.

STREET: 1755 S. Jefferson Davis Highway, Suite 400

CITY: Arlington

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961.716

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/283,917

FILING DATE: 03-AUG-1994

APPLICATION NUMBER: JP 209943/1993

FILING DATE: 03-AUG-1993

ATTORNEY/AGENT INFORMATION:

NAME: Oblon, No. 5880272man F.

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 2292-030-0

TELEPHONE: (703) 413-3000

TELEFAX: (703) 413-2220

TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 2095 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: cDNA

ORIGINAL SOURCE:

```

; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 844..2073
US-08-961-716-8

```

```

Query Match      2.4%; Score 50.8; DB 2; Length 2085;
Best Local Similarity 47.3%; Pred. No. 8.2e-06;
Matches 262; Conservative 0; Mismatches 277; Indels 15; Gaps 3;

```

```

QY 1001 TCTGCTTCCAGTATGATGAGAGAGTATCATTAACAGGATCATCGGATTTCCACGGTCAAG 1060
DB 1184 TCAATTTCCATCTGTTGTTTATGTTCTCTGCTTCAGAGGATGCTACAATTAAG 1243
QY 1061 TGTGGATGTAATACAGGTGAATGCTTAACACCTGATTACCATTTGTGAAGAGTTTC 1120
DB 1244 TGTGGGATATGAGATGAGAGATTTTGAACGAACTCTTAAGGGGCAATACAGACTCTGTAC 1303
QY 1121 TGCATTTGCGTTTCAATAATAG-----CATGATGGTGACCTGTCTCCAAAGATCGTTCCA 1174
DB 1304 AGGATATTTCAATTCGACCAAGTGGCAAGCTTCTGGCTTCATGTTCTGCAGATATGACCA 1363
QY 1175 TTGCTGTATGGGATATGGCTCCCACTGACATTAACCTCCGGAGGCTGCTGGTGGAC 1234
DB 1364 TTAAGCTATGGGATTTTCAGGCTTTGAATGATCAGAACCATGATGCGCATGAC---C 1420
QY 1235 ACCGAGCTGCTGTCAATGTTGTAGACTTTTGATGACAAAGTACATTTGTTCTGCATCTGGG 1294
DB 1421 ACAATGTTCTTCAGTAGCCATCATGCCCCAATGGAGATCAATATAGTGTCTGCCCAAGG 1480
QY 1295 ATAGAACTATAAAGTATGGAACACAGTACTTGTGAATTTGTAGGACCTTAATGGAC 1354
DB 1481 ATAAACTATAAAATGTGGGAAGTGCAAACTGGTACTGTGTGAAGACATTCACAGGAC 1540
QY 1355 ACAACGAGGCAATGCTGTTTGCAGTACAGGGA-----CAGGCTGGTAGTGTAGTGCT 1408
DB 1541 ACAGAGATGGGTACGTATGGTGGCCCAATCAAGACGGCACTCTGATAGCCAGCTGTT 1600
QY 1409 CATCTGACAACTATCAGATTTATGGGACATAGAAATGTTGGTGCATGTTTACAGAGTTAG 1468
DB 1601 CCAATGACCAAGCTGTGCGTGTATGGTGTAGCAACAAGGAATGCAAGGCTGAGCTTC 1660
QY 1469 AAGCCATGAGGAATGTTGCTGTTTGTATTCGATTTGATAACAAGAGGATAGTCAAGTGGG 1528
DB 1661 GAGAACATGACATGCTGTAGATGCAATTTCTGGGCTCCTGAAAGCTCATATCTTCCA 1720
QY 1529 CCTATGATGAAAA 1542
DB 1721 TCTCTGAAGCAACA 1734

```

RESULT 7

US-08-232-463-14/C

; Sequence 14, Application US/08232463

; Patent No. 5670367

; GENERAL INFORMATION:

; APPLICANT: DORNER, F.

; APPLICANT: SCHEIFLINGER, F.

; APPLICANT: FALKNER, F. G.

; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

; NUMBER OF SEQUENCES: 52

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 1800 Diagonal Road, Suite 500

; CITY: Alexandria

; STATE: VA

; COUNTRY: USA

; ZIP: 22313-0299

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25


```
/
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/232,463
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/07/935,313
/ FILING DATE:
/ APPLICATION NUMBER: EP 91 114 300.6
/ FILING DATE: 26-AUG-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BENT, Stephen A.
/ REGISTRATION NUMBER: 29,768
/ REFERENCE/DOCKET NUMBER: 30472/114 INMU
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703)836-9300
/ TELEFAX: (703)683-4109
/ TELEX: 899149
/ INFORMATION FOR SEQ ID NO: 14:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 7218 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ CLONE: pTZgpt-P1s
/
/ US-08-232-463-14
/
/ Query Match 2.3%; Score 50; DB 1; Length 7218;
/ Best Local Similarity 4.9%; Pred. No. 3.5e-05;
/ Matches 20; Conservative 218; Mismatches 168; Indels 0; Gaps 0;
/
/ QY 586 GAACCTGTGTGCAAGGAATGGTACCGAGTGACCTGTGATGGCATGCTGTGGAAGAGCTT 645
/   |||
/ Db 1448 GAAGAATTGGTACRBRBRBRBRBRBRBRBRBRBRBRBRBRBRBRBRBRBRBRBRBR 1389
/
/ QY 646 ATCGAGAGAAATGTCAGGACAGATCTCTGGAGAGCGCTGGCAGACAGAGAGATGG 705
/   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
/ Db 1388 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1329
/
/ QY 706 GGACAGCATATTTATCAAAACAACTCTGACGGGAATGCTCTCCCAACTTTTAT 765
/   : : : : : : : : : : : : : : : : : : : : : : : :
/ Db 1328 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1269
/
/ QY 766 AGAGCACTTTATCCCTAAATTAACAAGACATGTGAGCAATAGAACTTAATGGAGATG 825
/   : : : : : : : : : : : : : : : : : : : : : : : :
/ Db 1268 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1209
/
/ QY 826 GGAAGACATAGTTTACAGAGAAATTCACCTGCGAAGTGAACAAAGAGGAGTTACTGT 895
/   : : : : : : : : : : : : : : : : : : : : : : : :
/ Db 1208 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1149
/
/ QY 886 TTACAGTATGATCAGAAATAGTAAGCGGCTTCGAGACACACAAATCAAGATCTGG 945
/   : : : : : : : : : : : : : : : : : : : : : : : :
/ Db 1148 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1089
/
/ QY 946 GATAAAACACATGGAATGCAAGGCAATCTTCACAGGCCATACAG 991
/   : : : : : : : : : : : : : : : : : : : : : : : :
/ Db 1088 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1043
/
/ RESULT 8
/ US-08-188-582-17
/ Sequence 17, Application US/08188582
/ Patent No. 553410
/ GENERAL INFORMATION:
/ APPLICANT: Tjian, Robert
/ APPLICANT: Comai, Lucio
/ APPLICANT: Dynlacht, Brian D.
/ APPLICANT: Hoey, Timothy
/ APPLICANT: Ruppert, Siegfried
/ APPLICANT: Tanese, Naoko
/ APPLICANT: Wang, Edith
/ APPLICANT: Weinzierl, Robert O.J.
/
/
/ TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
/ TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
/ NUMBER OF SEQUENCES: 36
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
/ STREET: 4 Embarcadero Center, Suite 3400
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94111-4187
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/188,582
/ FILING DATE: 28-JAN-1994
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Osman, Richard A.
/ REGISTRATION NUMBER: 36,627
/ REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 781-1989
/ TELEFAX: (415) 398-3249
/ TELEX: 910 277299
/ INFORMATION FOR SEQ ID NO: 17:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2152 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 1..2112
/ US-08-188-582-17
/
/ Query Match 2.2%; Score 47.8; DB 1; Length 2152;
/ Best Local Similarity 52.6%; Pred. No. 8e-05;
/ Matches 131; Conservative 0; Mismatches 112; Indels 6; Gaps 1;
/
/ QY 1265 ATGCAAGTACATTTGTTCTGTCATCTGGGATAGAACTATAAAGGTATGGAACAAAGTA 1324
/   |||
/ Db 1625 ATTCCTAATTATGTTGCTACGGGCTCTCAGACAGAACTGTGCGGCTCTGGGAGCTCCTGA 1684
/
/ QY 1325 CTGTGAATTTGTAAGGACCTTAAATGGACACAAACGAGGCAATTCCTGTTGCAGTA-- 1382
/   |||
/ Db 1695 ATGTAACCTGTGTAAGGATCTTCACTGGACACAAAGGACCAATTCATTCCTTGACATTTT 1744
/
/ QY 1383 ----CAGGACAGGCTGGTAGTGCTGCTATCTGACAACTATCAGATTATGGGACA 1438
/   |||
/ Db 1745 CTCCTCAATGGGAGATTCCTGGCTACAGGACAAACAGATGGCAGAGTCTTCTTTGGGATA 1804
/
/ QY 1439 TAGAATGGTGTCATGTTTACGAGTGTAGAGGCAATGAGGCAATGGTGGGTTGTTTC 1498
/   |||
/ Db 1805 TTGACATGGTTGATGTTGGAGAAATTAAGGCACTGATACAGTCTGTTCACTTA 1864
/
/ QY 1499 GATTGATA 1507
/   |||
/ Db 1865 GGTTAGTA 1873
/
/ RESULT 9
/ US-08-646-715-17
/ Sequence 17, Application US/08646715
/ Patent No. 5637886
/ GENERAL INFORMATION:
/ APPLICANT: Tjian, Robert
/ APPLICANT: Comai, Lucio
/ APPLICANT: Dynlacht, Brian D.
/ APPLICANT: Hoey, Timothy
```



```

US-08-738-367-1
; Sequence 1, Application US/08738367
; Patent No. 5827688
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: Lavallie, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/738,367
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 433 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-738-367-1

Query Match 1.8%; Score 38.6; DB 1; Length 433;
Best Local Similarity 51.4%; Pred. No. 0.028; Indels 0; Gaps 0;
Matches 89; Conservative 0; Mismatches 84;

QY 1396 CTATGTAGTGCGCTCATCTGCACACACTATCAGATATTGGGACATAGAAATGGTGTGCATGT 1455
Db 107 GTCATAACTCTTCATGGGATAATAATGCTATTTTATTCATAGCATTTGGAGACGC 166
QY 1456 TTACGAGTGTTAGAGCCATGAGGAATGGTGCCTGTATTTCGATTTGATAACAAGAGG 1515
Db 167 CAGACACGGTATATGGGACATGATGCTGTATTAGATCTGTGTGCATGACAAAGG 226
QY 1516 ATATGTAGTGGGCGCTATGATGAAAAAATAAGTGTGGGATCTTGTGGCTGC 1568
Db 227 CTATATTCTGCTGGTGGACTCTACAGTGAAGGTGTGTCTGGTGTCTCTGC 279

RESULT 13
US-08-738-367-6
; Sequence 6, Application US/08738367
; Patent No. 5827688
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: Lavallie, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice

```

; APPLICANT: Spaulding, Vikki
 ; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
 ; TITLE OF INVENTION: ENCODING THEM
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genetics Institute, Inc.
 ; STREET: 87 CambridgePark Drive
 ; CITY: Cambridge
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02140
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/738,367
 ; FILING DATE:
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Brown, Scott A.
 ; REGISTRATION NUMBER: 32,724
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 498-8224
 ; TELEFAX: (617) 876-5851
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 733 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; US-08-738-367-5

Query Match 1.8%; Score 38.6; DB 1; Length 733;
 Best Local Similarity 51.4%; Pred. No. 0.04;
 Matches 89; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
 2Y 1396 GTAGTGGTGGCTCTCAGACACTATCAGATTATGGGACATAGATGGTGGCATGT 1455
 Db 107 GTCATAACTCTTCATGGGATAATATGCTATTTTATCCATAGCATTTGGAAGCG 166
 2Y 1456 TTACGAGTGTAGAGGCCATCAGGAATTGGTGGCTGTATTGCAATTGTAAACAAGG 1515
 Db 167 CAGGACACGTTAATGGGACATGATGCTGTTAGTAAGATCTGTGGCAAGACACAGG 226
 2Y 1516 ATAGTCAGTGGGCCCTATGATGGAATAATTAAGTGTGGATCTTGTGGCTGC 1568
 Db 227 CTATATCTGCATCGTGGGACTCTACAGTGAAGGTGCTGCTGTGCTGCTGCTG 279

RESULT 14
 US-09-156-425-1
 ; Sequence 1, Application US/09156425B
 ; Patent No. 5962671
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Brenda F.
 ; APPLICANT: Cowsett, Lex M.
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF PAN EXPRESSION
 ; FILE REFERENCE: RTS-0009
 ; CURRENT APPLICATION NUMBER: US/09/156,425B
 ; CURRENT FILING DATE: 1998-09-18
 ; NUMBER OF SEQ ID NOS: 47
 ; SEQ ID NO 1
 ; LENGTH: 3380
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (13)..(2766)
 ; US-09-156-425-1

Query Match 1.8%; Score 38.6; DB 2; Length 3380;
 Best Local Similarity 51.4%; Pred. No. 0.11;
 Matches 89; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
 QY 1396 GTAGTGGTGGCTCATCTGACAACTATCAGATTATGGACATAGAAATGTTGGTGCATGT 1455
 Db 2071 GTCATAACTCTTCATGGGATAATATGCTATTTTATCCATAGCATTTGGAAGCGC 2130
 QY 1456 TTACGAGTGTAGAGGCCATGAGGAATGGTGGCTGTATTGCAATTGTAACAAGAGG 1515
 Db 2131 CAGGACACGTTAATGGGACATGATGCTGTTAGTAAGATCTGTTGGCATGACACAGG 2190
 QY 1516 ATAGTCAGTGGGCCCTATGATGGAATAATTAAGTGTGGATCTTGTGGCTGC 1568
 Db 2191 CTATATCTGCATCGTGGGACTCTACAGTGAAGGTGCTGCTGTGCTGCTGCTG 2243

RESULT 15
 US-09-620-312D-1051
 ; Sequence 1051, Application US/09620312D
 ; Patent No. 6569662
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Chen, Rui-hong
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Yang, Yonghong
 ; APPLICANT: Wang, Jian-Rui
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Ma, Yungqing
 ; APPLICANT: Wang, Dunrui
 ; APPLICANT: Wang, Zhiwei
 ; APPLICANT: John Tillinghast
 ; APPLICANT: Drmanac, Radoje T.
 ; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
 ; FILE REFERENCE: Polypeptides
 ; FILE REFERENCE: 784CIP2B
 ; CURRENT APPLICATION NUMBER: US/09/620,312D
 ; PRIOR FILING DATE: 2000-07-19
 ; PRIOR APPLICATION NUMBER: 09/552,317
 ; PRIOR FILING DATE: 2000-04-25
 ; PRIOR APPLICATION NUMBER: 09/488,725
 ; PRIOR FILING DATE: 2000-01-21
 ; NUMBER OF SEQ ID NOS: 1105
 ; SOFTWARE: pt_FL_genes Version 1.0
 ; SEQ ID NO 1051
 ; LENGTH: 7518
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(7518)
 ; US-09-620-312D-1051

Query Match 1.7%; Score 36.8; DB 4; Length 7518;
 Best Local Similarity 53.5%; Pred. No. 0.75;
 Matches 77; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
 QY 1242 TGTGTCAATGTTGTAGACTTTGATGACCAAGTACATTTTCTGCACTCTGGGATAGAAC 1301
 Db 1563 TTCAGATATGGCAGTAACTATGAGAATACATGATGCTGCGGGAGCTGTGATAAAT 1622
 QY 1302 TATTAAGGTATGGAACACACAGTACTTGTGAATTTTAAGGACCTTAATGGACACACAG 1361
 Db 1623 TATTAGAGTGTGGTGTGTTGAGAACTTGTGCCCACTTGTGCTGCCAAGGACACACAGG 1682
 QY 1362 AGCAATGCTCTTTTGCAGTACAG 1385

db 1683 ATCAATTACATCTTACAGTTAG 1706

Search completed: October 22, 2003, 21:59:58
Job time : 150 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: October 22, 2003, 10:15:16 ; Search time 43 Seconds
(without alignments)

1272.558 Million cell updates/sec

Title: US-09-601-168B-2

Perfect score: 3034

Sequence: 1 MPAAEVLQKALKFNWSE.....PAAQAEPSPSRRTYVISR 569

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2597	85.6	518	2 B48088	beta-transducin re
2	1635.5	53.9	701	2 T16607	hypothetical prote
3	690	22.7	506	2 T50211	WD-repeat protein
4	590.5	19.5	605	2 T38932	probable sulfur me
5	545	18.0	640	2 S49932	MEF30 protein - ye
6	531.5	17.5	650	2 T46660	sulfur controller
7	520	17.1	579	2 T22703	hypothetical prote
8	519.5	17.1	267	2 S62507	hypothetical trp-a
9	455.5	15.0	1356	2 T18521	beta transducin-li
10	453	14.9	775	2 T45136	WD repeat protein
11	413.5	13.6	1227	2 AE1810	WD-40 repeat prote
12	399	13.2	779	2 S56245	cell division cont
13	396.5	13.1	703	2 T43557	F-box/WD-repeat pr
14	378.5	12.5	1189	2 A12493	WD-repeat protein
15	375	12.4	1747	2 AC1842	WD-40 repeat prote
16	374	12.3	1526	2 AC2239	WD-40 repeat prote
17	373.5	12.3	1258	2 A12155	WD-repeat protein
18	373.5	12.3	1683	2 AF2071	WD-40 repeat prote
19	361.5	11.9	677	2 AE1861	serine/threonine k
20	358.5	11.8	559	2 AB2502	hypothetical prote
21	354	11.7	409	2 S36113	Lis-1 protein - hu
22	354	11.7	410	2 S48052	platelet-activatin
23	341	11.2	515	2 S19487	hypothetical prote
24	339.5	11.2	777	2 T41075	hypothetical WD-re
25	337	11.1	1146	2 A55532	myosin-heavy-chain
26	336	11.1	676	2 AH2195	hypothetical prote
27	334	11.0	589	2 AG2400	WD-repeat protein
28	333.5	11.0	1711	2 AD1842	WD-40 repeat prote
29	332.5	11.0	317	2 T46032	WD-40 repeat regul

ALIGNMENTS

RESULT 1

B48088

beta-transducin repeat-containing protein - African clawed frog

N;Alternate names: beta-trcp

C;Species: Xenopus laevis (African clawed frog)

C;Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 21-Jul-2000

C;Accession: B48088

R;Spevak, W.; Keiper, B.D.; Stratowa, C.; Castanon, M.J.

Mol. Cell. Biol. 13, 4953-4966, 1993

A;Title: Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in anaphase are

A;Reference number: A48088; MUID:93330289; PMID:8393141

A;Accession: B48088

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-518 <SPE>

A;Cross-references: GB:M98269; NID:G295542; PIDN:AAA02810.1; PID:G295543

C;Superfamily: unassigned WD repeat proteins; WD repeat homology

C;Keywords: duplication

F;431-462/Domain: WD repeat homology <WD1>

Query Match 85.6%; Score 2597; DB 2; Length 518;

Best Local Similarity 91.6%; Pred. No. 1.2e-193;

Matches 488; Conservative 7; Mismatches 8; Indels 30; Gaps 1;

QY	18	SSERDCNNGEP	PRKIIPEKNSLRQTYNSCARLCINQETVCLASTAMKTENCVA	TKLAN 77
DB	13	ASERDCNRDEP	PRKIIITEKNTLRQ-----	TKLAN 42
QY	78	GTSSMIVPKQKLS	SYSEKELCVKYPQWSDQVFEVHLISQCHYOHGHSNYLK 137	
DB	43	GTSSMIVPKQKLS	ANYEKEKELCVKYPQWSECDQVFEVHLISRMCHYOHGHSNYLK 102	
QY	138	PMLODFITALP	ARGLDHIAENILSYLDLDAKCAELVCKEYRVTSDGMLWKKLIERMV 197	
DB	103	PMLODFITALP	ARGLDHIAENILSYLDLDAKCAELVCKEYRVTSDGMLWKKLIERMV 162	
QY	198	RTDSLWRGLAER	RGWQYLFKNKPPDGNAPNSFYRALYPKIIQDIETIESNWRGHSLS 257	
DB	163	RTDSLWRGLAER	RGWQYLFKNKPPDGTTPNSFYRALYPKIIQDIETIESNWRGHSLS 222	
QY	258	QRIHCRSETSK	GVYCLQYDDQKIYVSGLRDNTIKIWDKNTLECKRLTGTGSLCLOYDE 317	
DB	223	QRIHCRSETSK	GVYCLQYDDQKIYVSGLRDNTIKIWDKNTLECKRLTGTGSLCLOYDE 282	
QY	318	RVIIITGSSDS	TVRVWDVNTGEMLNTLIHHCBAVLHLRFNNGMVTCKSDRSIAVWDMASP 377	
DB	283	RVIIITGSSDS	TVRVWDVNTGEMLNTLIHHCBAVLHLRFNNGMVTCKSDRSIAVWDMASA 342	
QY	378	TDITLRRVLV	GHRAAVNVVDFDKYIVSASGDRTIKVNWTSTCFEFTLNHGKRGICLQ 437	
DB	343	TDITLRRVLV	GHRAAVNVVDFDKYIVSASGDRTIKVNWTSTCFEFTLNHGKRGICLQ 402	

QY 438 YRRLVVGSSDNTIRLWDIEGACLRVLVEGHEHVRIRPDKNKIVSGAYDGKIKWDL 497
 Db |||||
 QY 403 YRRLVVGSSDNTIRLWDIEGACLRVLVEGHEHVRIRPDKNKIVSGAYDGKIKWDL 462
 Db |||||

QY 498 VAALDRAPAGTLCUATLVHSGRVERLQDFDEFOIVSSSHDDTILWDFLNDP 550
 Db |||||
 QY 463 VAALDRAPAGTLCUATLVHSGRVERLQDFDEFOIVSSSHDDTILWDFLNDP 515
 Db |||||

RESULT 2

T16607
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T16607
 R:Miller, N.
 submitted to the EMBL Data Library, June 1995
 A:Description: The sequence of C. elegans cosmid K10B2.
 A:Reference number: Z18545
 A:Accession: T16607
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-701 <MOL>
 A:Cross-references: EMBL:U28730; NID:g860694; PID:g860695; PIDN:AAAG8258.1; CESP:K10B2.1
 A:Experimental source: strain Bristol N2
 C:Genetics:
 A:Introns: 78/3; 125/1; 183/2; 281/3; 404/3; 551/3; 668/3

Query Match 53.9%; Score 1635.5; DB 2; Length 701;
 Best Local Similarity 57.7%; Pred. No. 7,9e-119;
 Matches 326; Conservative 69; Mismatches 116; Indels 55; Gaps 8;

QY 31 RXIPEKNSURQYNSCARLQNOETVCLAS*AMKTCNCAKTKAN-----GTS 80
 Db |||||
 Db 2 RRFREGKALKRGARDGSGIOALVCVST----IERCF--TAVSNPIFFLFTFFSVF 55

QY 81 SMIVPKQR----KLSAYEKEKELCVKFEOWSSDOVERVEHLISOMCHYOCHINSY 135
 Db |||||
 Db 56 SPFPPSRNTQIFLUSGRSFSFSEVL----KMSHEQDFDMKIVHRLSHYQKQDNF 110

QY 136 LKPMQLQDFITAPARGLDHIAENILSYLDAKSLCAELVCKEYRVTS DGMKMKLIER 195
 Db |||||
 Db 111 IRPMLQDFISNLPA---HLVELILFNVNSDLSKCEVSTSWRCALARGHAKKLIK 166

QY 196 MVRTDSLWRGLAERGGGOVL-----FKNKPDPGNAPPNSFEAL 235
 Db |||||
 Db 167 NVKSDSLWGLSEKQMDKFNLSRMSVRRICEKFNVDNIRKDLQOLIMHVFYSKL 226

QY 236 YPKIIDIETIESNRCGRHSQRIRHCRSETSGVYCLQYDQKIVSGLRDNTIKINDKN 295
 Db |||||
 Db 227 YPKIIRDIHNDNNKRGYKMRINCOENGKGVYCLQYDDDKIVSGLRDNTIKINDRK 286

QY 296 TLECKRILTGTSVLCLOYDERVILITGSSDSTVRVMDVNTGEMLNTLIHCEAVLHLP 355
 Db |||||
 Db 287 DYSCRLSLSGHTGSVLCQYDNRVILITGSSDATVRVMDVNTGEMLNTLIHCEAVLHLP 346

QY 356 NNGMMVTCSDRSIAVMDASPTDITLRLVVGHRAAVNVDFDDKIVSASGDRTIKW 415
 Db |||||
 Db 347 ANGIWVTCSDRSIAVMDVMSPRDITIRRLVVGHRAAVNVDFDDKIVSASGDRTIKW 406

QY 416 NTSFCEVRLNGHKGRIACLOYDRVLVVGSSDNTIRLWDIEGACLRVLVEGHEHVR 475
 Db |||||
 Db 407 SMDTLEFVRTLAGHRRIGACLOVRGLVVGSSDNTIRLWDIEGACLRVLVEGHEHVR 466

QY 476 IRFDNKIVSGAYDGKIKWDLVAALDRAPAGTLCUATLVHSGRVERLQDFDEFOIVSS 535
 Db |||||
 Db 467 IRPDKIVSGAYDGKIKWDLQAALDRALSSIEICLSLVQHTGRVRLQDFDEFOIVSS 526

QY 536 SHDDTILWDFLNDPAAQAPPPSPRT 563
 Db |||||
 Db 527 SHDDTILWDFLNDP-----PSGLPSPRT 549

RESULT 3

T50211
 WD-repeat protein [imported] - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 02-Sep-2000
 C:Accession: T50211
 R:McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Brown, S.; Murphy, L.; Jones, L.; Mc:
 submitted to the EMBL Data Library, January 2000
 A:Reference number: Z25046
 A:Accession: T50211
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-506 <MOL>
 A:Cross-references: EMBL:AL136538; PIDN:CAB66464.1; GSPDB:GN00066; SPDB:SPAC30.05
 A:Experimental source: strain 972h(-); cosmid c30
 C:Genetics:
 A:Gene: SPAC29B6.01; SPDB:SPAC30.05
 A:Map position: 1
 A:Introns: 43/1; 74/3
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match 22.7%; Score 690; DB 2; Length 506;
 Best Local Similarity 30.4%; Pred. No. 1.3e-45;
 Matches 163; Conservative 104; Mismatches 167; Indels 102; Gaps 15;

QY 67 ENCVAK----TKLANGTSSMIVPKQRKLSAYEKEKELCVKFEOWSSDOVERVEHLIS 122
 Db |||||
 Db 8 KNVYSKVSDDLTCSDFSSTSPVPCNLPLS-----HENRIDLRDLA 50

QY 123 QMCHVOHQHINSYLPKMLQDFITAPARGLDHIAENILSYLDAKSLCAELVCKEYRV 182
 Db |||||
 Db 51 SLSEKGVVAVVNHVRSLLFTDFTVEFP-----EVSURVFSYLDQLDLCKCKMSRKWL 106

QY 183 TSDGMLWKLLI-----ERMVRTDSLWRG-----LAERRGWG----- 213
 Db |||||
 Db 107 LEDPGIKALYKQKGFVNENLNEPEARRTTKFPPOPREFNLKQNTIGPYGTWFLPQ 166

QY 214 QYLFNKPDPGNAPPNSFEALYKIIQDIETIESNRCGRHSQRIRHCRS----- 264
 Db |||||
 Db 167 QPFI-----DSNGRFLNWSYLY---KEHAHLDSDNNRGRFLVSTFNPFIRFPADQDF 217

QY 265 -ETSGVYCLQYDQKIVSGLRDNTIKINDKLTLECKRILTGTSVLCLOYDER--VII 321
 Db |||||
 Db 218 RATLDSVYCVQYDDEIMVSGSKORTVSVMDVNSRFILYKLYGHSGSVLCDFCERENLIV 277

QY 322 TGSSDSTVRVMDVNTGEMLNTLIHCEAVLHLPNMGMMVTCSDRSIAVW--DMASTPD 379
 Db |||||
 Db 278 SGSSDSTIIIDWQNRRLPKYVFGHTDNVLGVVSVSENYIISSRDHTARVRLDATSPAE 337

QY 380 ITRRLVVGHRAAVNVDFDDK--YIVSASGDRTIKWMTSTCEFVRTLNHGRGIACIQ 437
 Db |||||
 Db 338 ACM-HVLGHLASVNSVQYSSKTLIVTASSDRTLTWDITTHGCIIRIIHAGQGIACQ 396

QY 438 YDRVLVVGSSDNTIRLWDIEGACLRVLVEGHEHVRIRPDKNKIVSGAYDGKIKWDL 497
 Db |||||
 Db 397 YNGKEIVSGSSDLTIRIFBASGKLLMLQGHEDLIRTVRFNDEKIVSGYDGTVRIMW- 455

QY 498 VAALDRAPAGTLCUATLVH-----SGRVERLQDFDEFOIVSSSHDDTILWDF 546
 Db |||||
 Db 456 -----FNTGEQHCVLHNSRNSRVGLQDFDHRRIIACHTSHSEILWNWF 497

RESULT 4

T38932
 probable sulfur metabolite control protein - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 26-May-2000
 C:Accession: T38932
 R:Badcock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, April 1997
 A:Reference number: Z21818
 A:Accession: T38932

C;Function:
A;Description: negatively regulates sulfur structural gene expression
A;Note: scon-2+ expression is dependent on Cys3 function and the binding of Cys3 to the
C;Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match 17.5%; Score 531.5; DB 2; Length 650;
Best Local Similarity 25.0%; Pred. No. 3.6e-13;
Matches 144; Conservative 76; Mismatches 180; Indels 177; Gaps 11;

QY 117 VEHLISOMCHYQGHCHINSYKPMQLQDFITALPARGLDHIAENILSYLDAKSLCAELV 176
Db 99 LOGILSLQCFQPSFVSREVNEALKIDFISALPV-----ELAQVLCYLDTVSLTKAAQVS 154
QY 177 KEWYRVTSQDMLKKKLTIERMVRTDSLWRGLAERRGWQYLFKNKP-----P 222
Db 155 QRWRTLADSDAVWVRMCEQHVNK-----CTKCGWGLPFLERKKLRWYTRQOLAKGP 208
QY 223 DQNA-----PNSFYRA 234
Db 209 QGRVTELDADSHDSQDRSVNQHGRPAEAEEDPIKKRCMAAAEASKAVTPKTRSWKA 268
QY 235 LYPKIIQDIETIESNWRCCGRHSQRIHCRSETSKGYVCLQYDDQKIVGSLRDNTIKIWDK 294
Db 269 VY-----RDRQVSYNKNSEYKLSVL---KXGHEWVTCQLDDNLIATGSDYDTIKIWI 321
QY 295 NTLCKRILTGTGVLCLQYDERVITGSSDSTVRVMDVNTGEMNLTIHHCEAVLHLR 354
Db 322 ETECIRTLYGHTAGIRALQFDSDSKLISGLDHTIKWNMTHTGECSTFAAHTDSVISVH 381
QY 355 FNGMVMVTCSDRSIAVMDMASPTDITLRRVLVGRRAAVN--VDFDDKIYVSAGDRTI 412
Db 382 FDGHLLASGSDTKVTFDNSEKTYCLK-----GHSDWNSTHVDIKSTVFSASDDTTI 437
QY 413 KVNTSTCEVRLTNGHKGRI-----ACLQYRDLVVS----- 445
Db 438 KLADLDTRQIRTYEGHVGHVQVQLPPEYEPDEVLNCGASQDNQDAMSYSGGSGSPS 497
QY 446 -----GSSD 449
Db 498 MSHAQIERAGSPGSHSSHNLLPSSLPSGDEVDHLYGSFAFVADRSRPLPRYFNTGSLD 557
QY 450 NTRILWDIEGACLRVLGHEHELVRCIRFDNKRIVSGAVDGKIKVMDLVAALDPRAPAGT 509
Db 558 STWRLLWDSARGLRLTLFGHLEGWGLAGDTIRVIRSGANDGMVKTWE-----PRSG 608
QY 510 LCLRTLVEHSGRVRFLQFQFQVSSSHDDTILWDF 546
Db 609 KCDATYTGCGPVCVGLSDSLMASGSDGTIRLHSF 645

RESULT 7
T22703
A;Description: hypothetical protein F55B12.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T22703
R;Sims, M.
submitted to the EMBL Data Library, September 1996
A;Reference number: Z19602
A;Accession: T22703
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-579 <WIL>
A;Cross-references: EMBL:Z79757; PIDN: CAB02129.1; GSPDB: GN00023; CESP: F55B12.3
A;Experimental source: Clone F55B12
C;Genetics:
A;Gene: CESP: F55B12.3
A;Map position: 5
A;Introns: 22/3; 45/2; 77/3; 122/3; 171/2; 211/2; 342/2; 417/3; 478/3; 526/1

Query Match 17.1%; Score 519.5; DB 2; Length 267;
Best Local Similarity 38.0%; Pred. No. 8.8e-33;
Matches 104; Conservative 58; Mismatches 83; Indels 29; Gaps 6;

QY 286 DNTIKIMDKNILECKRILTGTGTVSCLQYDER--VITGSSDSTVRVMDVNTGEMNL 343
Db 1 DRTVSVQVNVNRFILYKLYHGSGVSLCLDFCRRLNLVSGSSDSTIIIDWQNRPLKVY 60
QY 344 IHCCEAVLHFRFNGMVMVTCSDRSIAVW--DNASPTDITLRRVLVGRRAAVNVDFFDK 401
Db 61 FGHDTNVLGVVSVSENYIISSSRDHTARVRLDATSPAEACM--HVLRGHLASVNSVQYSK 119
QY 402 --YIVSAGSDRTIKWNVTSTCEFVRTLNGHKGRIACLOYRDLRVVSGSSDNTIRLWDIEC 459
Db 120 TGLIVTASSDRTLRTWDTITGHGCIIRIIHAHQRIACQYNGKFIYVSGSSDLTIRIFEAS 179

QY 72 KTKLANGTSSMI-----VPKORK--LSAYEKEKEL-----CVKYFEQMSDQVE 115
Db 27 ESSYNGSSSYNADKLSSSRPLQHKLDLSPSRNDLNPVRVHILIALFKDLSSAEQMD 86
QY 116 FVEHLISOMCHYQGHCHINSYKPMQLQDFITALPARGLDHIAENILSYLDAKSLCAELV 175
Db 87 AFTRLLQBSNNNTNIRQLRAIEHPQDFLSCLPV-----ELGMKILHNLGTVDLLKVAQV 142
QY 176 CKWYRVTSQDMLKKKL--IERMV-----RTDSLWRGLAERRGWQYLFKNKPPDGNAP 227
Db 143 SKNWKLISEIDKIWKSILGVBEFKHPDPDTRVTGAMQGTAAAG-----VTIPDHIOQ 195
QY 228 PN-SFYRALYPKIIQDI-----ETESNWRCCGRHSQRIHCRSETSKGYVCL 273
Db 196 COLNVHRLFKLOKFGDIFERAADSKRVLRADKTEKNWNPIMGSAV--LRGHEDHUITCM 254
QY 274 QYDDQKIVSGURDNTIKIWDKNTLECKRILTGTGTVSCLQYDE--RVITGSSDSTVRV 331
Db 255 QIHDDVLVTGSDNTLXVWCIDKGEVMTLVGTGGVTSQISQCGRYIVSGSTORTVKV 314
QY 332 WDVAVTGEMNLTIHHCEAVLHFRFNGMVMVTCSDRSIAVMDMASPTDITLRRVLVGRHA 391
Db 315 WSTDVGSLLHILQGTSTVRCAMAGSILVTGSDTYLRVWDVSEGRHLA---TLGHHA 371
QY 392 AVNVVDPDKIYVSAGDRTIKWNVTSTCEFVRTLNGHKGRIACLOYRDR--LVVSGSSD 449
Db 372 AVRCYQFQDGTTVVSGGYDFTVKIWNHTGRCIRTLGHNNRVYSLLFESERSIVCSGSLD 431
QY 450 NTRILWDI---ECGACLRVLGHEHELVRCIRFDNKRIVSGAYDGKIKVMDLVAALDPRAP 506
Db 432 TSIRVWDITRREGQECVALLQGTSLTSGWQLRGNLVSCNADSHVRVWDI-----H 483
QY 507 ACTLCRLTVLHSGRVRFLQ--FDEFQIVSSSHDDTILWID 545
Db 484 EGT-CVHMLSGHRAITSLQWFGERNMVATSSDGTVKLWD 522

RESULT 8
S62507
hypothetical trp-asp repeat-containing protein - fission yeast (Schizosaccharomyces pombe
C;Species: Schizosaccharomyces pombe
C;Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
C;Accession: T38502; S62507
R;Jones, L.; Murphy, L.; McNeil, A.; Simpson, I.; Harris, D.; Barrell, B.G.; Rajandream.
submitted to the EMBL Data Library, October 1995
A;Reference number: Z21798
A;Accession: T38502
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-267 <JO2>
A;Cross-references: EMBL:Z66525; NID:G1044926; PIDN:CAA91423.1; PID:G1044927; GSPDB:GN00
C;Genetics:
A;Map position: 1
A;Superfamily: unassigned WD repeat proteins; WD repeat homology
P;18-51/Domain: WD repeat homology <WDL>
F;104-137/Domain: WD repeat homology <WD2>

Query Match 17.1%; Score 519.5; DB 2; Length 267;
Best Local Similarity 38.0%; Pred. No. 8.8e-33;
Matches 104; Conservative 58; Mismatches 83; Indels 29; Gaps 6;

QY 286 DNTIKIMDKNILECKRILTGTGTVSCLQYDER--VITGSSDSTVRVMDVNTGEMNL 343
Db 1 DRTVSVQVNVNRFILYKLYHGSGVSLCLDFCRRLNLVSGSSDSTIIIDWQNRPLKVY 60
QY 344 IHCCEAVLHFRFNGMVMVTCSDRSIAVW--DNASPTDITLRRVLVGRRAAVNVDFFDK 401
Db 61 FGHDTNVLGVVSVSENYIISSSRDHTARVRLDATSPAEACM--HVLRGHLASVNSVQYSK 119
QY 402 --YIVSAGSDRTIKWNVTSTCEFVRTLNGHKGRIACLOYRDLRVVSGSSDNTIRLWDIEC 459
Db 120 TGLIVTASSDRTLRTWDTITGHGCIIRIIHAHQRIACQYNGKFIYVSGSSDLTIRIFEAS 179

Qy 460 GACLRVLEHGEHLVRCIRPFNKIVSGAYDGKIKVWDLVAALDPRAPAGTCLRLTLVBH- 518
Db 180 GKLRMLQGHEDLIRTVRFNEIVSGDVGTVIRWN-----FNTGEQHC 224
Qy 519 -----SGRYERLQFDEFQIVSSSHDITILIMDF 546
Db 225 VLHNSRNRVFGLOFDRHRIIACHSSEILVWNF 258

RESULT 9
T18521
beta transducin-like protein - Podospora anserina
C:Species: Podospora anserina
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18521
R:Saupé, S.; Turcq, B.; Begueret, J.
Gene 162, 135-139, 1995
A:Title: A gene responsible for vegetative incompatibility in the fungus Podospora anserina
A:Reference number: Z18944; PMID:96009891; PMID:7557402
A:Accession: T18521
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1356 <SAU>
A:Cross-references: EMBL:L28125; NID:g607002; PID:g607003; PIDN:AAA85775.1
C:Genetics:
A:Gene: het-el
A:Introns: 761/3

Query Match 15.0%; Score 455.5; DB 2; Length 1356;
Best Local Similarity 31.7%; Pred. No. 8.2e-27;
Matches 121; Conservative 64; Mismatches 131; Indels 66; Gaps 15;
Qy 226 APPNSFYRALY---PKIIQDIETIESNRCGRHSL-----QRHCRS- 264
Db 802 APTDSMIKKIKKEPGWISTISWEAEWNACTQTLEHGSSVLSVAFSADQORVASGD 861
Qy 265 -----ETSKG-----VTCLOY--DDQKIVSGLRDNTIKWKNLTCEKRL 303
Db 862 DTKIKINDTASGTQTLEHGSSVMSVAFSPRRERVASGSDTKIKINDAASGCTOTL 921
Qy 304 TGTGTVLCLOY--DERVIITGSSDSTVRVMDVNTGEMLNTLIHCEAVLHLEN-NGMM 360
Db 922 EHGGRVQSVAFSPDQORVASGSDHTIKINDAASGCTQTLEHGSSVLSVAFSPDQOR 981
Qy 361 VTC-SKDRSTAVMDASPTDITLRVLVGHRAVNVVDF--DDKIYVSASGRTIKWNT 417
Db 982 VASGSGDKTIKINDTASG--TCTQTLLEGHGGVMSVAFSPDQORVASGSDTKIKINDT 1038
Qy 418 STCEFRVTLNGHKGIACTOYR--DRLVVSAGSDNTIRLWDIECGACLRVLEHGEHLVRC 475
Db 1039 ASGTCTQTLLEGHGGVMSVAFSPDQORVASGSDHTIKINDAVSGTCTQTLLEGHGSVWS 1098
Qy 476 IRF--DNKRIVSGAYDGKIKVWDLVAALDPRAPAGTCLRLTLVEHSGRVRIQF--DEFQ 531
Db 1099 VAFSPDQORVASGSDTKIKIND-----AASGT-CTQTLLEGHGGVMSVAFSPDQOR 1149
Qy 532 IVSSSHDITILIMDFLNDPAAQ 553
Db 1150 VASGSDTKIKINDAASGCTQ 1171

RESULT 10
T45136
WD repeat protein popi [imported] - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000
C:Accession: T45136; T40157
R:Kominami, K.; Toda, T.
submitted to the EMBL Data Library, September 1996
A:Description: Fission yeast WD repeat protein Popi is involved in maintenance of ploidy
A:Reference number: 222925
A:Accession: T45136

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-775 <KOM>
A:Cross-references: EMBL:Y08391; PIDN:CAA69671.1
A:Experimental source: strain h- 972
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Devlin, K.; Churcher, C.M.
submitted to the EMBL Data Library, March 1998
A:Reference number: Z21842
A:Accession: T40157
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-775 <WOO>
A:Cross-references: EMBL:AL022103; PIDN:CAA17898.1; GSPDB:GN00067; SPDB:SPB2C2G2.18
A:Experimental source: strain 972h; cosmid c2G2
C:Genetics: SPEC2G2.18
A:Gene: SPEC2G2.18
A:Map position: 2
A:Note: pop1+

Query Match 14.9%; Score 453; DB 2; Length 775;
Best Local Similarity 25.8%; Pred. No. 5.8e-27;
Matches 162; Conservative 86; Mismatches 229; Indels 150; Gaps 20;
Qy 9 QEKALKFMNSSEEDC-----NNGEPFRKIPEKNSLRQTY-NSCARLCLNQET 56
Db 121 REKCLKRRNSSLNLHANKRFLFNSQSDGNKKQNETPEFTNYSNVFYPNNCDKSEVASET 180
Qy 57 VC-----LASTAMKTENCVAKTKLANGTSSMIVPKORKLSA----- 92
Db 181 TFSLDAPNNSVNYVSFPNLLGNDKTKRQSPFSSSSSHNSLHEPVIYDFSSNPISHP 240
Qy 93 ----SYEKEKEL----CVKYFEQWSESDQVFVHLSIQMCHYQHGHINSYKPMLODF 144
Db 241 SNHLSQKNAVLKLAQLISSFEKLPESVRQYLLFELLSSRCGHAVQNIHKILLPIFOKNF 300
Qy 145 ITALPARGLDHTAENILSYLDAKSLCAELACKWYR-VTSDGMLWKKL----- 193
Db 301 LTGFPA-----ETNLVLTHLDAPSLCAVSQVSHHWIKLVSSNEELWKSFLKOGFFWDSI 356
Qy 194 ERMVRTDLSMRGLAERRGQYLFKNKPPDGNAPPNSFYRALYPKIIQDIETIESNRCG 253
Db 357 DSKIRTMCLEQLSA-----CAIMKRVYFRHFNLRERW--- 389
Qy 254 RSLQRIHCRSETSKGYVC---LOYDOKIVSGLRDNTIKWKNLTCEKRLTGHGTSV 310
Db 390 LHAPEKIKCSPPIHGVRILITKLQFDDDKIIVSTCSPRINTYDTKTGVLKSLSEHEDV 449
Qy 311 LCLQYDERVVIITGSSDSTVRVMDVNTGEMLNTLIHCEAVLHLENMVMVTCSDKRSIA 370
Db 450 WTEYVVDTLVTGSDTTRVWDLTGECKQVFGTSTIRCIKIVGQNSTTDTD--- 505
Qy 371 VMDMASPTDITLRVLVGHRAVNVVDFDDKIVVSASGRTIKWNT----- 416
Db 506 -----DVEKE-----NRPASNDANSMPPIYSSSRDCTIRLWSLPCLDPPFPVNVNE 552
Qy 417 -----TSTCEFRVTLNGHK---RGIACLOYRDLRVVSGSSONTIRLWDIECGAC 462
Db 553 NPDQNDFTSATNPPYIRTLRGHTDSVREACL---GDLIVSASYDGTLLRVKASGTVC 609
Qy 463 LRVLEGEHELVRICIRPDNKR--IVSGAYDGKIKVWDLVAALDPRAPAGTCLRLTLVEHSG 520
Db 610 LHVLRGHVGRVSVTINFSRQOCISAGTDAKIRIWNL-----ESGEL-LQTLHGHSN 660
Qy 521 RVPRLOQDEFQIVSSS--HDDTILIMD 545
Db 661 LVSQVTFNQNILVSASAPDTSIRVWD 687

RESULT 11
AE1810
WD-40 repeat protein [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

C;Accession: AE1810
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE1810
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-1227 <KUR>
A;Cross-references: GB:BA000019; PIDN:BA077553.1; PID:gi717135007; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr0029

Query Match 13.6%; Score 413.5; DB 2; Length 1227;
Best Local Similarity 23.9%; Pred. No. 1.3e-23;
Matches 140; Conservative 114; Mismatches 208; Indels 125; Gaps 22;
Qy 10 EKALKFMNSGERE-----DNNNGEPPRKIIPEKNSLRQTYNSCARLCLNQETVCLA 60
Db 378 EKKVMYSLGANREYVSPRELKDDWLTATESPIKMEALESLLR-----RSLIEKASPTLI 431
Qy 61 STAMKTENCVAKTKLANGTSMIVPKORLKSASYEKELVKYFROESDOVEFVEHL 120
Db 432 EKASTQGEKEAESKFGLESVWN---EYITAKFIENS-----DEEFSQTKLDF---- 478
Qy 121 ISOMCHYQHGHSYLYKPMQLQDFITAPALPARGLDHIAENILSYLDAKSLCAELVCKEY 180
Db 479 -----INTY--PLMK-----ARSLDYI----- 493
Qy 181 RVTSQGLMKKLIBRMVTSLSLARGLAERGWQYLFKPKNPPDGNAPPN--SFYALYPK 238
Db 494 ROIQERILIEPVKOLNIFGTLEHLRLMGLTLOKEPLPKKGVAAGNLNLLRQLQOLD 553
Qy 239 IIPDIETIESN-----WRCGRHSL---QRHCRS-----ETSKGVYCLQY--D 276
Db 554 KIPDESPIILSGRDFGLTIQWAFYKEVKLEKIFANSDLTSGVFTETMSSVSVKFSFD 613
Qy 277 DQKIVSGLRDNTIKWPKNTLECKRIILTHGTGSLVCLQY--DERVITSSSDSTVRWDV 334
Db 614 GKVFATGLMNGEIRLWQTSNQLRYKGTAWVAFSPDSRMLASGASDSTIKLMDV 673
Qy 335 NTGEMLNTLIHCEAVLHRAF--NNGMMVTCSDRGSIATVMDWASPTDITLRVLVGHRAA 392
Db 674 HTGCLXTLSKNTKNSVAFSPDGRILASASQDTIKLWDIATG--NCQOTLIGHDDW 730
Qy 393 VNVVDF---DDK--YIVSASGRTIKWNTSTCEFVRTLNHKGRIACLQYR--DRLVV 444
Db 731 VWSVTFSPVTDREPLLLASSADQHIKMDVATGKCLKTLKGTREHVSFSPDQTLA 790
Qy 445 SGSSDNTIRLWDIECGACLVLEGHEBELVRCIRF--DNKRIVSGAYDGKIKVMDVLAALD 502
Db 791 SSGSDSTVRLWDVKTGQWQIFEGHSKKVSVRFSPDGTQLASCGEDRSIKLWDIQRG-- 848
Qy 503 PRAPAGTILCLRTLVHSGRVRLOF--DEQIVSSSHDDTILWDPL 547
Db 849 -----ECVNTLWGHSSQWMAIFSPDGRITLISCSDDQTLARLWDVI 888

RESULT 12

S56245
Cell division control protein CDC4 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YFL009W
C;Species: Saccharomyces cerevisiae
C;Date: 02-Sep-1995 #sequence_revision 12-Apr-1996 #text_change 26-May-2000
C;Accession: S56245; S48310; A26867; S62304
R;Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sasano, submitted to the EMBL Data Library, May 1995
A;Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces cerevisiae
A;Reference number: S56186
A;Accession: S56245

A;Molecule type: DNA
A;Residues: 1-779 <MUR>
A;Cross-references: EMBL:D50617; NID:9836685; PIDN:BA009229.1; PID:dl009870; PID:g836745
R;Churcher, C.
submitted to the EMBL Data Library, September 1994
A;Reference number: S48310
A;Accession: S48310
A;Molecule type: DNA
A;Residues: 1-579 <CHU>
A;Cross-references: EMBL:Z46255; NID:9559925; PIDN:CAA86341.1; PID:9559926; MIPS:YFL009W
R;Yochem, J.; Byers, B.
J. Mol. Biol. 195, 233-245, 1987
A;Title: Structural comparison of the yeast cell division cycle gene CDC4 and a related
A;Reference number: A26867; MUID:88011240; PMID:3309335
A;Accession: A26867
A;Molecule type: DNA
A;Residues: 1-459, 'E' 461-779 <YOC>
A;Cross-references: EMBL:X05625; NID:93502; PIDN:CAA29113.1; PID:g3503
R;Murakami, Y.
submitted to the EMBL Data Library, April 1994
A;Reference number: S62302
A;Accession: S62304
A;Molecule type: DNA
A;Residues: 1-779 <MUM>
A;Cross-references: EMBL:D31600; NID:9836814; PIDN:BA006495.1; PID:dl007066; PID:g836815
C;Genetics:

A;Gene: SGD:CDC4
A;Cross-references: SGD:S0001895; MIPS:YFL009W
A;Map position: 6L
C;Function:
A;Description: initiation of DNA replication; separation of the spindle pole bodies to form a new spindle pole body
C;Superfamily: unassigned WD repeat proteins; WD repeat homology
C;Keywords: cell cycle control
F;459-494/Domain: WD repeat homology <WD1>
F;628-659/Domain: WD repeat homology <WD2>

Query Match 13.2%; Score 399; DB 2; Length 779;
Best Local Similarity 24.3%; Pred. No. 9.1e-23;
Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

Qy 56 TVCLASTAKTENCVAKTKLANGTSMIVPKORLKSASYEK-----ELCVKVFQW 108
Db 194 TPLAKTTKINN-----NNNIADLIESKDSITSPYLSDEIFSAINNKLPHAYFK-- 244
Qy 109 SPSQDOVEFVEHLISOMCHYQHGHSYLYKPMQLQDFITAPALPARGLDHIAENILSYLDAS 168
Db 245 -----NLLFRLVANWDRSELSDGLTIKDNLKRDLITSLP-----EISLKIENYLOFED 294
Qy 169 LCAAEVLVCKEYR-VTSDGMLWKKLI--ERMVRTDSLWRGLAERGWQYLFKPKNPPDGN 225
Db 295 IINSLGVSONNKKIIRKSTSLWKKLISENFV-----SPKGF 331
Qy 226 APPNSFYALYPKIQQD-----IET--IESNWRGCRHSLQRHCRSETSKGVYCLQYD 276
Db 332 NSLNKLKSQYFKLSQDRLRLSLFLENIFLKNWVNFVQRTTLRGHMTSVITCLQFE 391
Qy 277 DQKIVSGLRDNTIKWPKNTLECKRIILTHGTGSLVCLQYDE-RVITIGSSDSTVRWDVN 335
Db 392 DNYVITGADDKMIRVYDSINKKFLQLSGHDGGVWALKYAHGGILVSGSDTDTVRWMDIK 451
Qy 336 TGEMLNTLIHCEAVLHRAF--NNGMMVTCSDRGSIATVMDWASPTDITLRVLVGHRAAVN 395
Db 452 KG-----CCT-----HVFKGHNSTVRC 468
Qy 396 VPDD---KYIVSASGRTIKWNT-----STCE-----FVRTLNG 428
Db 469 LQIVYKNIKVIYTGSRDNTLHVKLPKESVPDGHBEHDYFLVFHTPEENPYFVGLRG 528
Qy 429 HKEGACLVLRDLVYSSGSSDNTILMDIEGACLVLEGHEBELVRCIRPDN--KRVISG 486
Db 529 HMASVTVSGHGNIVVSGSYONTLLVMDVAMKCLYLISGHTDRIYSTIYDHEKRCISA 589
Qy 487 AYDGKIKWDL-----VAALDPAPAGTL--CLRTLVEHSGRVRLOFDEQFQIVSS 535

Db 589 SMDTIRIWDLENIMNNGECYATNSGPCAKILGAMYTLQHTALVGLRLSDKFLVSA 648
QY 536 SHDDTILINDFLNDPAQAQBPSPRTVY 566
Db 649 AADGSRGWD-AND-----YRKFSY 668

RESULT 13
T43557
F-box/WD-repeat protein pop2 - fission yeast (Schizosaccharomyces pombe)
N:Alternate names: proteolysis factor sud1p
C:Species: Schizosaccharomyces pombe
C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C:Accession: T43557; T38794; T43798
R:Wolf, D.A.; Jackson, P.K.
A:Description: Fission yeast pop2 encodes a novel F-box/WD-repeat protein involved in th
A:Reference number: 222576
A:Accession: T43557
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-703 <MOL>
A:Cross-references: EMBL:AF038867; PIDN:AAB95480.1
R:Gentiles, S.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1996
A:Reference number: 221812
A:Accession: T38794
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-703 <GEN>
A:Cross-references: EMBL:Z98602; PIDN:CAB11275.1; GSPDB:GN00066; SPDB:SPAC4D7.03
A:Experimental source: strain 972h-; cosmid C4D7
R:Jallepalli, P.V.; Tien, D.; Kelly, T.J.
Proc. Natl. Acad. Sci. U.S.A. 95, 8159-8164, 1998
A:Title: Sud1+ targets cyclin-dependent kinase-phosphorylated Cdc18 and Runt proteins fo
A:Reference number: 222686; MUID:98318628; PMID:9653157
A:Accession: T43798
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-703 <JAL>
A:Cross-references: EMBL:AF064515; NID:G3293382; PIDN:AAC39496.1; PID:G3293383
C:Genetics:
A:Gene: pop2; sud1; SPAC4D7.03
A:Map position: 1
C:Function:
A:Description: required to prevent spontaneous re-replication

Query Match 13.1%; Score 396.5; DB 2; Length 703;
Best Local Similarity 23.3%; Pred. No. 1.2e-22;
Matches 146; Conservative 92; Mismatches 226; Indels 163; Gaps 25;
QY 14 KFMSSRED--CNGGPPPKI-----IPKNSLRQYNSCARLCL-----NOETVCL 59
Db 96 KFNEIENVSKCLDHPSPDPVPGDSISLVPQPNFL-YSHSLPPKIISIDNRNRIKL 154
QY 60 ASAMTKENCVAKTKLANGTSSMIVPKQKLSASYEK-EKELCVKYFPGWSESDQVEFVE 118
Db 155 NSISSNSDNFPFPPKV--DTSNTVSPGSPISDELDLNLQISVOTFEDLPEGIQ--SYAF 211
QY 119 HLISQMCHYQGHG-INSYKPMQLQRFITAPARGLDIAENILSYLDAKSLCAAEVLCK 177
Db 212 FQLRSNCQSMRLLLNECEPLKXILSNLFP-----SIVQILLNLDHSLFLSP 267
QY 178 EYRVV---TSDGMLWKKLIERMVRTDSLWRGUAERGGQVLFKXNPPDGNAPP-----N 229
Db 268 TWRNILDVHTS---YWKHMF-----SLFGQINENDW-----KYANENLRNPPFLHND 312
QY 230 SFYALYPKLIQDITETTESNRCGRHSLOP-----
Db 313 QISDDYFEIFK-----RHFLNRKRWLFPSPSHLSPPFIHVPNFMITSLLLH 360

QY 260 -----IHRSETSK-GVYCLQYDDQXIVSGLRDNTIKIWKNT 296
Db 361 KDRITITSGSGTIOHNAITGVLEARLEGHKEGVAVKIHENTLVSSGIDKTVRVWNIK 420
QY 297 LECKRILTGHGTSGVLCQY-----DERVLIITGSSDSTVRVMDVNTGEML 340
Db 421 AKCTHIFRGHLSIIRCLIELVPSRLIRHGVEIPEPDQYIVSGSRDHTLRVW----- 472
QY 341 NTLIHCEAVLHFRNNGMMVTCSDRSIAVMDMASPTDITLRVLVGHRAAVNVVDFDD 400
Db 473 -----KLPKNTDPPYLPDNTNSIDRWE-KNDFYFV---HTLIGHTDSVRTISGYG 517
QY 401 KYIVSASGDRITIKYWNSTCEFFVTLNGHKGACLOYRD--RLVVGSSDNTIRLWDIE 458
Db 518 DVLVSGSDSIRIRWVSTGECLYHLRGHSRLYVLYFERNICSGSMKIRVWDLS 577
QY 459 CGACLRVLEGHEELVRCIRFDNKRIVSGAYDGKIKWDLVAALDPRAPAGTLCRLTVEH 518
Db 578 TCTCKVYLEGHDAPVTLNVPQNLISGSADSTIRWD---LNTGKP---LWVLPNS 628
QY 519 SCRVRLOQDFEQIVSSSHDDTILWD 545
Db 629 SCYISSFVSDBKII-SCNDGSKVLWD 654

RESULT 14
AI2493
WD-repeat protein [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120alpha
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AI2493
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
DNA Res. 8, 205-213, 2001
A:Title: Complete genome sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AI2493
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1189 <KUR>
A:Cross-references: GB:BA000020; PIDN:BA078213.1; PID:gl7135667; GSPDB:GN00180
C:Genetics:
A:Gene: alr7129
A:Genome: plasmid

Query Match 12.5%; Score 378.5; DB 2; Length 1189;
Best Local Similarity 26.7%; Pred. No. 6.4e-21;
Matches 96; Conservative 63; Mismatches 107; Indels 93; Gaps 10;
QY 280 IVSGLRDNTIKIWKNTLECKRILTGHGTSGVLCQYD--ERVITGSDSTVRVMDVNTG 337
Db 745 LASGFDQNVKLWDIHTGKVTMLQGTGVVTSVAFNPKDNLGSGYDQSVKWDKRTG 804
QY 338 EMLNTLIHCEAVLHFRN-----
Db 805 RCLDTLKXHTNRINSVAFHPQCHLFSGGDDHAAKIHELGTGQCICKTQGHSNATYIIAH 864
QY 357 ---NGMMVTCSDRSIAVMD--MASPTDITLR-----RLVVGHRAAVNVVDFDD--KYIVS 405
Db 865 NWEHSLLAGSHEDQTIKLDNLNLHSPKSNVNTHPFRILQHSNRVFSVFSSTGQLLAS 924
QY 406 ASGDTIKVWNTSTCEFFVTLNGHKGACLOYR--DLVVGSSDNTIRLWDIECCACL 463
Db 925 GSADRTIKLWSPHTQCCLHTLHGGSWWNAFASLDDKLLASGYDHTVKIWDVSSQCL 984
QY 464 RVLGHEELVRCIRF--DNKRIVSGAYDGKIKWDL-----VAALDP 503
Db 985 QTLQHGPGSLAVASCQKTLFSSGYEKLQKQWDETGYCLOQWEADSNRVAVSRD 1044
QY 504 RAPAGT-----LCRLTVEHSGRVRLOP--DEFQIVSSSHDDTILWD 545

Db 1045 NOYLATGDDSVVRLMDIGKGVCTRFSGHTSOVICILFTKDCRRWISSSDRIKIWN 1103

RESULT 15

AC1842
WD-40 repeat protein [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AC1842
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AC1842
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1747 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA077807.1; PID:G17135261; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all0283

	Query Match	12.4%	Score 375;	DB 2;	Length 1747;
	Best Local Similarity	32.7%	Pred. No. 2.1e-20;	Mismatches 107;	Indels 16; Gaps 7;
Matches	84;	Conservative	50;		
Qy	276	DDQKIVSGLRDNTIKIWDKNTLECKRILTGHTGSLVLCLOY--DERVIITGSSDSTVRVWD	333		
Db	1487	DGKTIATASADNTIKLWDSQTQQLIKTLTGKDKRITLTSFHPDNQTIASGSADKTIKIWR	1546		
Qy	334	VNTGEMLNTLIHCEAVLHLRF--NNGMMVTCSDRSIAVMDASPTDITLRRVLVGHRA	391		
Db	1547	VNDGQLRLTGTGNDVTSVNFSPDQGLASGSTNTVKIWQ----TDGRLIKNTIGHL	1602		
Qy	392	AVNVVDF--DDKXIVASGDRTIKWNTSTCEFVRTLNHGKRGIAQLQYR--DRLVWSGS	447		
Db	1603	AJASVKFSPDSHTLASASWDNTIKLWQVTDGKLNNKLNHIDGVTLSLSPSPDGEILASGS	1662		
Qy	448	SDNTIRLWDIECGACLVLEGHEELVRCIRF--DNKRIVSGAYDGKIKVWDLVAALDPR	505		
Db	1663	ADNTIKLWNPATLLKTLTGHPGKINTLAFSPDGKTLSSGGEDAGVWVWNL--DLDDL	1720		
Qy	506	PAGTLCRLTIVHSGRV	522		
Db	1721	QQGCDRITDYLQHSNV	1737		

Search completed: October 22, 2003, 10:26:03
Job time : 48 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

1M protein - protein search, using sw model

run on: October 22, 2003, 10:07:11 ; Search time 26 Seconds

(without alignments)

1029.161 Million cell updates/sec

Title: US-09-601-168B-2

Perfect score: 3034

Sequence: 1 MDPAAVLQEKALKFMSSE.....PAAQAEPRSPRTYTYISR 569

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3006	99.1	605	FW1A HUMAN	Q9V297 homo sapien
2	2597	85.6	518	TRCE XENLA	Q91854 xenopus lae
3	2384.5	78.6	542	FW1B HUMAN	Q9UKB1 homo sapien
4	1638.5	54.0	665	L123 CAEEL	Q09990 caenorhabdi
5	690	22.7	506	POFB SCHPO	Q09855 schizosacch
6	590.5	19.5	605	POF1 SCHPO	P87053 schizosacch
7	575	19.0	678	SCOB EMENI	Q00659 emericella
8	545	18.0	640	MT30 YEAST	P39014 saccharomyc
9	531.5	17.5	650	SCO2 NEUCR	Q01277 neurospora
10	520	17.1	579	SE10 CAEEL	Q91794 caenorhabdi
11	477.5	15.7	684	CDG3 CANAL	P53699 candida alb
12	455.5	15.0	1356	HET1 PODAN	Q00808 podospora a
13	453	14.9	775	POPI SCHPO	P87060 schizosacch
14	399	13.2	779	CDG4 YEAST	P07834 saccharomyc
15	396.5	13.1	703	POP2 SCHPO	O14170 schizosacch
16	392	12.9	732	KMHB DICDI	P90648 dictyosteli
17	374	12.3	1526	YY46 ANASP	Q8YR11 anabaena sp
18	373.5	12.3	1258	YS00 ANASP	Q8YTC2 anabaena sp
19	373.5	12.3	1683	YL24 HUMAN	Q8YV57 anabaena sp
20	354	11.7	409	L1S1 HUMAN	P43034 homo sapien
21	354	11.7	409	L1S1 MOUSE	P43035 mus musculu
22	353	11.6	409	L1S1 BOVIN	P43033 bos taurus
23	341	11.2	515	YK24 YEAST	P25382 saccharomyc
24	337	11.1	1146	KMWA DICDI	P42527 dictyosteli
25	334.5	11.0	422	FW2 HUMAN	Q9UKT8 homo sapien
26	325.5	10.7	361	WDS DROME	Q9V3J8 drosophila
27	324.5	10.7	334	WDR3 HUMAN	Q9UGP9 homo sapien
28	318	10.5	422	FW2 MOUSE	P60584 mus musculu
29	313.5	10.3	376	YKY4 CAEEL	Q17963 caenorhabdi
30	312.5	10.3	714	YJL2 YEAST	P47025 saccharomyc
31	307.5	10.1	742	PKWA THECU	P49695 thermomonos
32	307.5	10.1	1693	Y163 SYNY3	Q55563 synecocyst
33	306	10.1	704	T2D4 DROME	P49846 drosophila

34	238.5	9.8	800	1	T2D4 HUMAN	Q15542 homo sapien
35	294.5	9.7	1249	1	APAF RAT	Q9EPV5 rattus norv
36	292	9.6	659	1	YK16 YEAST	P36130 saccharomyc
37	290.5	9.6	614	1	TU11 SCHPO	Q09715 schizosacch
38	289	9.5	1249	1	APAF MOUSE	O88679 mus musculu
39	287.5	9.5	327	1	GBUP BRANA	Q39336 brassica na
40	287	9.5	1248	1	APAF HUMAN	O14727 homo sapien
41	286.5	9.4	327	1	GBUF ARATH	O24456 arabidopsis
42	285.5	9.4	798	1	T2D4 YEAST	P38129 saccharomyc
43	283	9.3	473	1	PRP5 SCHPO	O13615 schizosacch
44	283	9.3	713	1	TU11 YEAST	P16649 saccharomyc
45	281.5	9.3	582	1	TU11 KLULA	P56094 kluyveromyc

ALIGNMENTS

RESULT 1

FW1A HUMAN	STANDARD;	PRT;	605 AA.
AC	Q9V297; Q9V213;		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	F-box/WD-repeat protein 1A (F-box and WD-repeats protein beta-TrCP)		
DE	(E3RSIKappaB) (pikappaBalpha-E3 receptor subunit).		
GN	BTRC OR FBXW1A OR FBXW1A OR BTRCP.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A. (ISOFORM 1).		
RX	MEDLINE=99075339; PubMed=9859996;		
RA	Yaron A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,		
RA	Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.;		
RT	"Identification of the receptor component of the IkappaBalpha-		
RT	ubiquitin ligase."		
RL	Nature 396:559-564 (1998).		
RN	[2]		
RP	SEQUENCE FROM N.A. (ISOFORM 2).		
RC	TISSUE=Lymphoid;		
RX	MEDLINE=98323370; PubMed=9660940;		
RA	Margottin F., Bour S.P., Durand H., Selig L., Benichou S., Richard V.,		
RA	Thomas D., Strebel K., Benarous R.;		
RT	"A novel human WD protein, h-beta TrCP, that interacts with HIV-1 Vpu		
RT	connects CD4 to the ER degradation pathway through an F-box motif."		
RL	Mol. Cell 1:565-574 (1998).		
RN	[3]		
RP	SEQUENCE FROM N.A. (ISOFORM 2).		
RX	MEDLINE=20003060; PubMed=10531035;		
RA	Cenciarelli C., Chaur D.S., Guardavaccaro D., Parks W., Vidal M.,		
RA	Pagano M.;		
RT	"Identification of a family of human F-box proteins."		
RL	Curr. Biol. 9:1177-1179 (1999).		
RN	[4]		
RP	CHARACTERIZATION.		
RX	MEDLINE=99145464; PubMed=9990852;		
RA	Winston J.T., Strack P., Beer-Romero P., Chu C.Y., Ellledge S.J.,		
RA	Harber J.W.,		
RT	"The SCF(beta-TRCP)-ubiquitin ligase complex associates specifically		
RT	with phosphorylated destruction motifs in I-kappa-B-alpha and		
RT	beta-catenin and stimulates I-kappa-B-alpha ubiquitination in vitro."		
RL	Genes Dev. 13:1270-1283 (1999).		
RN	[5]		
RP	INTERACTION WITH PHOSPHORYLATED CTNNB1.		
RX	MEDLINE=2072105; PubMed=12077367;		
RA	Sadot E., Conacci-Sorrell M., Zhurinsky J., Shnizer D., Lendo Z.,		
RA	Zharhary D., Kam Z., Ben-Ze'ev A., Geiger B.;		
RT	"Regulation of S33/S37 phosphorylated beta-catenin in normal and		
RT	transformed cells."		
RL	J. Cell. Sci. 115:2771-2780 (2002).		
CC	!- FUNCTION: Substrate-recognition component of the SCF (SKP1-CUL1-F-		


```

box protein) ubiquitin ligase complex, which mediates the
ubiquitination of proteins involved in cell cycle progression,
signal transduction and transcription. Regulates the stability of
CTNBN1 and participates in wnt signaling.
-!- PATHWAY: Ubiquitin conjugation; third step.
-!- SUBUNIT: Interacts directly with SKP1 in the SCF complex.
Interacts specifically with phosphorylated CTNBN1 and NFKB1A,
ubiquitination substrates.
-!- SUBCELLULAR LOCATION: Cytoplasmic.
-!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=1;
IsoId=Q9Y297-1; Sequences=Displayed;
Name=2;
IsoId=Q9Y297-2; Sequences=VSP 006764;
-!- SIMILARITY: Contains 1 P-box domain.
-!- SIMILARITY: Contains 7 WD repeats.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AF101784; AAD08702.1; -.
EMBL; Y14153; CAA74572.1; -.
EMBL; AF129530; AAF04464.1; -.
Genew; HGNC:1144; BTRC.
MIM; 603482; -.
GO; GO:0004840; F-ubiquitin conjugating enzyme activity; TAS.
GO; GO:0007165; P-signal transduction; TAS.
GO; GO:0004511; P-ubiquitin-dependent protein catabolism; TAS.
InterPro; IPR001810; F-box.
InterPro; IPR001680; WD40.
Pfam; PF00646; F-box; 1.
Pfam; PF00400; WD40; 7.
PRINTS; PR00320; GPROTEINBRPT.
ProDom; PD000018; WD40; 4.
SMART; SM00256; FBOX; 1.
SMART; SM00320; WD40; 7.
PROSITE; PS00181; FBOX; 1.
PROSITE; PS00678; WD_REPEATS_1; 6.
PROSITE; PS00682; WD_REPEATS_2; 7.
PROSITE; PS00294; WD_REPEATS_REGION; 1.
UniProt; UniProtKB; Ligase; Wnt signaling pathway; Repeat;
WD repeat; Alternative splicing.
DOMAIN; 190 228
REPEAT; 301 338 WD 1.
REPEAT; 341 378 WD 2.
REPEAT; 381 418 WD 3.
REPEAT; 424 461 WD 4.
REPEAT; 464 503 WD 5.
REPEAT; 505 541 WD 6.
REPEAT; 553 590 WD 7.
VARSPLIC; 17 52 Missing (in isoform 2).
/FTIG-VSP 006764.
SEQUENCE 605 AA; 6886 MW; 4C67F3B7E400FD37 CRC64;
Query Match 99.1%; Score 3006; DB 1; Length 605;
Best Local Similarity 94.0%; Pred. No. 8 6e-220;
Matches 569; Conservative 0; Mismatches 0; Indels 36; Gaps 1;
2y 1 MDPAEVLQKALKFM-----NSEEEDC 24
|||||
3b 1 MDPAEVLQKALKFMPSRLWGLCSLADMPSLRCLNYPGTGALTAFONSEEEDC 60
|||||
2y 25 NCEPPRKIPKNSLRQYNCSARCLNQETVCLASTAMKTENCVAKTLANGTSMIV 84
|||||
3b 61 NCEPPRKIPKNSLRQYNCSARCLNQETVCLASTAMKTENCVAKTLANGTSMIV 120
|||||
2y 85 PKQKLSASYEKEKELCVKFEQWSESDQVEFVEHLISQMHYQGHINSYLPKPLQDF 144

```

```

|||||
121 PKQKLSASYEKEKELCVKFEQWSESDQVEFVEHLISQMHYQGHINSYLPKPLQDF 180
|||||
145 ITALPARGLHIAENILSYLDAKSLCAELVCKEWYRTSDGMLWKLIERVMTDSLWR 204
|||||
181 ITALPARGLHIAENILSYLDAKSLCAELVCKEWYRTSDGMLWKLIERVMTDSLWR 240
|||||
205 GLAERRGWGYLPKNKPPDGNAPNSFYRALYPKIIQDIETTESNWRGHSLOIHCRS 264
|||||
241 GLAERRGWGYLPKNKPPDGNAPNSFYRALYPKIIQDIETTESNWRGHSLOIHCRS 300
|||||
265 ETSKGVYCLQYDDQKIVSGLRDNTIKWDKNTLECKRILTHTGSLVCLQYDERVITGS 324
|||||
301 ETSKGVYCLQYDDQKIVSGLRDNTIKWDKNTLECKRILTHTGSLVCLQYDERVITGS 360
|||||
325 SDSVTRVMDVNTGEMNTLIHCEAVLHFRNNGMVTCKDRIAVMDMSPDTILRR 384
|||||
361 SDSVTRVMDVNTGEMNTLIHCEAVLHFRNNGMVTCKDRIAVMDMSPDTILRR 420
|||||
385 VLVGHRAAVNVDFDDKIIVSASGDRTIKVMNTSTCEFVRTLNHKGRIACIQYRDLVV 444
|||||
421 VLVGHRAAVNVDFDDKIIVSASGDRTIKVMNTSTCEFVRTLNHKGRIACIQYRDLVV 480
|||||
445 SGSSDNTIRLWDIECGACLRVLEGHELVRCIEFONKRIVSGAYDGIKVDLVAALDPR 504
|||||
481 SGSSDNTIRLWDIECGACLRVLEGHELVRCIEFONKRIVSGAYDGIKVDLVAALDPR 540
|||||
505 APAGTLCRLTVEHSGRVFRLQDFQIVSSHDDTILWDFLNDPAAQAEPPSPRTY 564
|||||
541 APAGTLCRLTVEHSGRVFRLQDFQIVSSHDDTILWDFLNDPAAQAEPPSPRTY 600
|||||
565 TYISR 569
|||||
601 TYISR 605
|||||
RESULT 2
TRCB XENLA
ID TRCB XENLA STANDARD; PRT; 518 AA.
AC Q91854; P70037; P70038;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Beta-TrCP (Beta-transducin repeat-containing protein).
GN FBXW1 OR BTRCP.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=93330289; PubMed=8393141;
RA Spevak W., Keiper B.D., Stratowa C., Castanon M.J.;
RT "Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in
anaphase are rescued by Xenopus cDNAs encoding N-ras or a protein
with beta-transducin repeats.";
RL Mol. Cell. Biol. 13:4953-4966(1993).
RN [2]
SEQUENCE OF 302-518 FROM N.A.
RP MEDLINE=97109804; PubMed=8952061;
RA Hudson J.W., Alarcon V.B., Elinson R.P.;
RT "Identification of new localized RNAs in the Xenopus oocyte by
differential display PCR.";
RL Dev. Genet. 19:190-198(1996).
CC -!- FUNCTION: Probably recognizes and binds to some phosphorylated
proteins and promotes their ubiquitination and degradation. May
participate in wnt signaling.
CC -!- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX
(BY SIMILARITY).
CC -!- DEVELOPMENTAL STAGE: PRESENT IN FULLY GROWN AND PROGESTERONE-
MATURED OOCYTES. THE LEVEL CHANGE VERY LITTLE EVEN AFTER ZYGOTIC
GENE TRANSCRIPTION BEGINS FOLLOWING THE MIDBLASTULA TRANSITION. DO

```



```

JR PRINTS; PRO0320; GPROTEINBRPT.
JR ProDom; PD000018; WD40; 4.
JR SMART; SM00256; FBOX; 1.
JR SMART; SM00320; WD40; 7.
JR PROSITE; PS00181; FBOX; 1.
JR PROSITE; PS00678; WD REPEATS 1; 5.
JR PROSITE; PS00682; WD REPEATS 2; 7.
JR PROSITE; PS00294; WD REPEATS REGION; 1.
CW Developmental protein; Cell cycle; Cell division;
CW Ub1 conjugation pathway; Repeat; WD repeat.
JT DOMAIN 81 127
JT REPEAT 220 257 WD 1.
JT REPEAT 260 299 WD 2.
JT REPEAT 301 337 WD 3.
JT REPEAT 343 380 WD 4.
JT REPEAT 383 420 WD 5.
JT REPEAT 423 460 WD 6.
JT REPEAT 472 509 WD 7.
JT MUTAGEN 441 441 G->R; IN LIN-32 (RH293).
JT SEQUENCE 665 AA; 75916 MW; BF3E9AF5F12ECC CRC64;
3Q

Query Match 54.0%; Score 1638.5; DB 1; Length 665;
Best Local Similarity 63.7%; Pred. No. 2.6e-116;
Matches 316; Conservative 59; Mismatches 92; Indels 29; Gaps 3;

2Y 88 RKLASAYEKEKELCVKPEOMSEDOVEFVHLISOMCHYOHGHNSVLYKPLQDFITA 147
2b 27 KPLSDYLOGHGLIEVLKSEHQDPMDKIVHRLSHYQLGKVDNFIKPLQDFISN 86
2Y 148 LPARGLDHIAEILSYLDKASICAALVCKEYRVYTSQDMLWKXLIEMVYVTSLSMGLA 207
2b 87 LPA---HLVELILEFVNDSLSKCEEVSTWRCALARGQHWKGLIEKVRSDSLMWGLS 142
2Y 208 ERGKGQYL-----FNKPPDGNAPNSFVALYKPIODIETIE 247
2b 143 EKRWQDKFLNISRDMSVRIACEKNFYDNVNIKDKLDQLILMHVYKLYPIKIRIHND 202
2Y 248 SNMRCGRHSLQRIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIKDNTLBECKRLTGT 307
2b 203 NNWKGNYKWTIRINQSENSKGVYCLQYDDQKIVSGLRDNTIKIKDNTLBECKRLTGT 262
2Y 308 GSVLCLOYDERVLIIGSSDSYRVWDVNTGEMNTLIHCEAVLHLFNNGMVTCSDR 367
2b 263 GSVLCLOYDNVRIIGSSDSYRVWDVNTGEMNTLIHCEAVLHLFNNGMVTCSDR 322
2Y 368 STAVNDMASFTDITRRVLVGHRAAVNVVDFDKYIVSASGDRITKVMNTSTCFEVRTLN 427
2b 323 STAVNDMVSFDTITRRVLVGHRAAVNVVDFDKYIVSASGDRITKVMNTSTCFEVRTLN 382
2Y 428 GHRGTAQLOYRDLRVVSGSSDNTIRLWDIEGACLRVLEGHELVRCIRFDKRIYVGA 487
2b 383 GHRGTAQLOYRDLRVVSGSSDNTIRLWDIEGACLRVLEGHELVRCIRFDKRIYVGA 442
2Y 488 YDGKIKVMDLVAALDPRAPAGTLCRLTIRVHSGRVRFLQDFEFOIVSSSHDDTLIWDFL 547
2b 443 YDGKIKVMDLVAALDPRALSSICLSLVQHTGRVRLQDFDFOIVSSSHDDTLIWDFL 502
2Y 548 NDPAAQAEPRPSRT 563
2b 503 DAP-----PSGLPSST 513

RESULT 5
POFB_SCHPO STANDARD; PRT; 506 AA.
ID Q09855; Q9P7V1;
ID 01-FEB-1996 (Rel. 33, Created)
ID 16-OCT-2001 (Rel. 40, Last sequence update)
ID 28-FEB-2003 (Rel. 41, Last annotation update)
ID F-box/WD-repeat protein pof11.
ID POF11 OR SPAC239E.01 OR SPAC30.05.
ID Schizosaccharomyces pombe (fission yeast).

```

Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 Schizosaccharomycetales; Schizosaccharomycetaceae;
 Schizosaccharomycetes.
 NCBI_TaxID=4896;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21668955; PubMed=11809834;
 RA Katayama S., Kitamura K., Lehman A., Nikaido O., Toda T.;
 "Fission yeast F-box protein Pof3 is required for genome integrity and
 telomere function.";
 RT Mol. Biol. Cell 13:211-224(2002).
 RL [2]
 SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 Howarth S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 Skellern J., Simmonds M., Squares R., Stevens K.,
 Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 Woodward J., Vackaert G., Aert R., Robben J., Grymonprez B.,
 Welljens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
 Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forburg S.L.,
 Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
 "The genome sequence of Schizosaccharomyces pombe.";
 Nature 415:871-880(2002).
 RL Nature 415:871-880(2002).
 CC -!- SIMILARITY: Contains 1 F-box domain.
 CC -!- SIMILARITY: Contains 7 WD repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL; AB061694; BAB55543.1; -;
 DR EMBL; AL136538; CAB66464.1; -;
 DR EMBL; Z66525; CAA91423.1; -;
 DR PIR; T50211; T50211
 DR GeneDB_Spombe; SPAC239E.01; -;
 DR InterPro; IPR001810; F-box.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00646; F-box; 1.
 DR Pfam; PF00400; WD40; 7.
 DR PRINTS; PRO0320; GPROTEINBRPT.
 DR ProDom; PD000018; WD40; 1.
 DR SMART; SM00256; FBOX; 1.
 DR SMART; SM00320; WD40; 8.
 DR PROSITE; PS00181; FBOX; 1.
 DR PROSITE; PS00678; WD REPEATS 1; 3.
 DR PROSITE; PS00682; WD REPEATS 2; 6.
 DR PROSITE; PS00294; WD REPEATS REGION; 1.
 KW Repeat; WD repeat.
 FT DOMAIN 70 116 F-BOX.
 FT REPEAT 219 256 WD 1.

```

2Y REPEAT 259 298 WD 2.
PT REPEAT 301 338 WD 3.
PT REPEAT 345 386 WD 4.
PT REPEAT 388 426 WD 5.
PT REPEAT 427 464 WD 6.
PT REPEAT 468 505 WD 7.
SQ SEQUENCE 506 AA; 58257 MW; CEP34DAEFTBC2E10 CRC64;

Query Match
Best Local Similarity 30.4%; Pred. No. 1e-44; Length 506;
Matches 163; Conservative 104; Mismatches 167; Indels 102; Gaps 15;

67 ENCVAK-----TKLANGTSMIVPKOKLSASVEKEKELCVKVFQWSSRSDQVFEVHLIS 122
8 KNVVSKVSDLTSCDSFSTSPVCLNPLS-----HENNRIDLIIRDLA 50

123 QMCHYQGHINSYLVKPLMLQDFITALPARGLHIAENILSYLDAKSLCAELCKEYWRV 182
51 SLSKEGVAVYNHVRSLLFDFTEVEP-----EVSRLVFSYLDQLDCKCKLMSKRWKRL 106

183 TSDGMLWKLI-----ERMVRTDSLWRG-----LAERRGWS----- 213
107 LEDEGIWKALYMKQKGFVNENVLNBEAMRTHKFPQPFENFLKQNMIGPYGTMLPQ 166

214 QYLFQKPPGPNAPPNSFYRALYPKIIQDIETIESNRCGRHSLORIHCRS----- 264
167 QFIF-----DSNGRPLNWSYLY---KEHAHLDSNWRHGRFLVSTFNNPISRFPAQDF 217

265 -ETSGKGYCLOYDQKIVSGRLNTIKIWDKNTLECKRILTGTGSLVLCLOYDER--VII 321
218 RATLDSYVCQYDDEIMWSSGKDRTVSVDNRSRFLYKLYGSGSVLCFCRRRLV 277

322 TGSSDTVRVADVNTGEMLNTLIHCEAVLHLRFNNGMVTCKDRSTAVW--DMASTD 379
278 SGSSDSTIIINDQNRRLPKLYVFGHTDNLGVVSENYIISSRDHTARVRLDATSPAE 337

380 ITRRVLVGHRAAVNVVDFDK--YIVSASGDRTIKWNTSTCEVRTLNGKHKGIAQLQ 437
338 ACM-HVLRGHLASVNSVYSSKGLIVTASSDRTLTWDTITGHGIRIHAHQGIACAQ 396

438 YRRLVVGSSDNTIIRLWDIEQACLRVLGEHBLVRCIRFPNKRIVSGVADGKIKWDL 497
397 YNGKFIYVSGSDLTIRIFEASSGKLLRLMQHEDLIRTVRENDKIVSGGYDGVIRWN- 455

498 VAALDPRAPAGTLCRLTVEH-----SGRVRLQFDFQVSSSHDDTILWDF 546
456 -----FNTGECHVCLHNSRNSRVFGLQFDHRIIACHTSHSEILVWNF 497

RESULT 6
ID POFI SCHPO STANDARD; PRT; 605 AA.
AC P87053;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE F-box/WD-repeat protein pof1 (Skp1-binding protein 1).
GN POFI OR SBP1 OR SPAC57A10.05C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
DC Schizosaccharomycetales; Schizosaccharomycetaceae;
CC Schizosaccharomycetes.
CX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=972;
RA Katayama S., Kitamura K., Lehmann A., Nikaido O., Toda T.;
RT "Fission yeast F-box protein Pof1 is required for genome integrity and
telomere function.";
RL Mol. Biol. Cell 13:211-224 (2002).
RN [2]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds R., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Art R., Robben J., Grynopre B.,
RA Weljens I., Vantreels E., Kieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Fumelle B.,
RA Coffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880 (2002).
CC -!- FUNCTION: Probably recognizes and binds to some phosphorylated
CC proteins and promotes their ubiquitination and degradation.
CC -!- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX
CC (BY SIMILARITY).
CC -!- SIMILARITY: Contains 1 F-box domain.
CC -!- SIMILARITY: Contains 7 WD repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB032410; BAA84528.1; --
DR EMBL; Z94864; CAB08168.1; --
DR PIR; T38932; T38932.
DR GeneDB; Spombe; SPAC57A10.05c; --
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001880; WD40.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINBRPT.
DR ProDom; PD000018; WD40; 3.
DR SMART; SM00255; FBOX; 1.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS50181; FBOX; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 2.
DR PROSITE; PS50082; WD_REPEATS_2; 7.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Ub1 conjugation; Repeat; WD repeat.
FT DOMAIN 107 153 F-BOX.
FT REPEAT 271 299 WD 1.
FT REPEAT 311 339 WD 2.
FT REPEAT 350 379 WD 3.
FT REPEAT 390 420 WD 4.
FT REPEAT 432 460 WD 5.
FT REPEAT 472 500 WD 6.
FT REPEAT 510 538 WD 7.
SQ SEQUENCE 605 AA; 67110 MW; 7118C9379EC5C1F0 CRC64;

Query Match
Best Local Similarity 19.5%; Score 590.5; DB 1; Length 605;
Matches 156; Conservative 86; Mismatches 189; Indels 107; Gaps 12;

```

QY 80 SSMIVPQKRLSASYPEKELCVKPEQWSE---SDQVEFVEHLISQMCHYOHGHSYLL 136
 Db 46 SSM-----HNELSGLSEKQRVEAVMAAFSEACSEKRLAQGLNNSSLLSFASSTL 101
 QY 137 KPMQORDFITALPARGLDHIAENILSYLDKSLCAELVCKEYRVVTSQGLMVKKLIERM 196
 Db 102 DSLVRLDFLSLPV-----SISFRILSFUDARSLCQAQVSKHKELADDDVWHRMCEQH 157
 QY 197 VRTDLSLWGLAERRGQVLFKN----- 219
 Db 158 INRK-----CEKCGWGLPFLERNTLYAAKASIQRYERLTKRGVDQAHESPPVKKAKLD 211
 QY 220 -----KPPDGNAPENSP-----YRALYKLIQDIETIESNWRGRHSLO 258
 Db 212 DYPSSNEETISSVKPPSPNSKFPKLPKTPKVEVAERCR-----VECNWRHGR----- 263
 QY 259 RIHCR-----SETSGVYCLQYDDQKIVSGLRDNTIKIMDRNTLECKRILTGTGTVLCLQ 314
 Db 264 ---CQWLSSHSDGVMCLVRNLSAGSYDATIRLNLATFQVQVALLGHSSTGVTCLQ 320
 QY 315 YDERVITGSSDSTVRVWNTGEMLNTLIHCEAVILHLRPNNGMVTCSKDRSIADVDM 374
 Db 321 FQCKLISSMDKIRIMWYRSECSILHGTDSVLCITFDSTLULVSGSADCTVKLWHF 380
 QY 375 ASPTDITLRRVLVGHRAAVNVVDF--DDKYIVSASGDRITIKWNTSTCFVRTLNGHKG 432
 Db 381 SGGKRITLR---GHTGPNVRIIRDRGLVLSGSDSTIKIWSLENTNLTCLHFSAHIGP 436
 QY 433 IACQVRDLVSGSSDNTIRLWDECGACLRVLGHEBELVRCIRPDKRIVSGAYDGI 492
 Db 437 VQSLALADSLFSCSLDGTIKQWDIEKKCVHTLFGHIEGVMEIAADHLRLISGHDGV 496
 QY 493 KWDVLAALDPRAPACTLCLRTLVHSGRVFLQDFEFOIVSSSHDDTILWDFLNDP 550
 Db 497 KWWEACE-----CVHTLKHSEPTVVALGDCVSGSDGKIYLWLFNNAP 543

RESULT 7
 SCOB_EMENI STANDARD; PRT; 678 AA.
 AC Q00659;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Sulfur metabolite repression control protein.
 GN SCOB OR MAPB1.
 OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; Emericella.
 OX NCBI_TaxID=162425;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Natoriff R.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: REGULATORY PROTEIN THAT CONTROL SULFUR METABOLITE
 CC REPRESSION.
 CC -!- SIMILARITY: Contains 8 WD repeats.
 CC -!- SIMILARITY: Contains 1 F-box domain.
 CC -!- SIMILARITY: BELONGS TO THE MET30/SCOB/SCON-2 FAMILY OF WD-REPEAT
 CC PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC ENBL; U21220; AAC15905.1; --
 DR InterPro; IPR001810; F-box.
 DR InterPro; IPR001680; WD40.

DR Pfam; PF00646; F-box; 1.
 DR Pfam; PF00400; WD40; 7.
 DR PRINTS; PR00320; GPROTEINBRPT.
 DR ProDom; PD000018; WD40; 3.
 DR SMART; SM00256; FBOX; 1.
 DR SMART; SM00320; WD40; 7.
 DR PROSITE; PS50181; FBOX; 1.
 DR PROSITE; PS00678; WD_REPEATS_1; 4.
 DR PROSITE; PS00882; WD_REPEATS_2; 7.
 DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
 KW Transcription regulation; Repeat; WD repeat.
 FT DOMAIN 178 224 F-BOX.
 FT REPEAT 347 375 WD 1.
 FT REPEAT 387 415 WD 2.
 FT REPEAT 427 455 WD 3.
 FT REPEAT 466 496 WD 4.
 FT REPEAT 508 543 WD 5.
 FT REPEAT 553 595 WD 6.
 FT REPEAT 607 635 WD 7.
 FT REPEAT 647 675 WD 8.
 SQ SEQUENCE 678 AA; 76070 MW; D840D452E37B4C53 CRC64;

Query Match 19.0%; Score 575; DB 1; Length 678;
 Best Local Similarity 26.9%; Pred. No. 7.4e-36;
 Matches 149; Conservative 83; Mismatches 168; Indels 154; Gaps 14;

QY 117 VEHLSIQMCHYGHINSYLPMLQORDFITALPARGLDHIAENILSYLDKSLCAELVLC 176
 Db 153 LOGIWAQCCFPOLSVISATVRLIRIDFITALP-----PEIAFKILCYLDTTSLCKASQVS 208
 QY 177 KPWYRVTSQGLMVKKLIERMVVRTDSLRGLAERRGQVLFKNK----- 220
 Db 209 RGWRALADDDVVMHRCQHIHRK-----CKKCGWGLPDLDRKELRESKREIELRATTW 262
 QY 221 -----PPDNA--PP-----NSFYRLY---P 237
 Db 263 DKGVGPSPDASAEPSPPSGKKKJEDDEVAVVXKHSLSGSDAGVDKSDFFKTRYPWK 322
 QY 238 KIQDIETIESNWRGRHSRQRIHCRSETSK---GVYCLQYDDQKIVSGLRDNTIKIW 293
 Db 323 EYVKDRFKVGTWNYGR-----CSIKTFKHTNGVMCLQFEDNLTATGSDYDTTIKW 375
 QY 294 KNTLECKRILTGTGTVLCLQYDERVITGSSDSTVRVWNTGEMLNTLIHCEAVLHL 353
 Db 376 TTGTEELRLTGHESGIRCLQFDQDQTKLISGSMRDTIKVMWRTGECISTYTGHRGGVIGL 435
 QY 354 RFNNGMVTCSKDRSIADVDMASPTDITLRRVLVGHRAAVNV--VFDDKYIVSASGDR 411
 Db 436 HFDAISILASGVDTKVKIWNFEKSTPFSLR-----GHTDWNANRVDTSSRTVFSASDDCT 491
 QY 412 IKVMTSTCFVRTLNGH-----KRGIA----- 435
 Db 492 VRLWDLTKTCIRTPHGHVGVQVQVPLPREFEFEEHDAECENDDLTSTSGDANPPSIQA 551
 QY 436 -----LQYR-----RLVVGSSDNTIRLWDIECGACLRVLEHLEL 472
 Db 552 SMGLEPNAAYSQSSAFGTSFNGRAAPRYVWTSALDSTIRLWTTTGRCLRTFEGHLEG 611
 QY 473 VRCIRPDKRIVSGAYDGIKVKVWDLVAALDPRAPAGTLCLFTLVHSGRVFLQDFEFOI 532
 Db 612 VWALGADTLRIVSGAEDRMIKIW-----PRTGKCERTFTGHSQPVTCIGLGSRF 662
 QY 533 VSSSHDDTILWDF 546
 Db 663 ATGSEDCVEVRMYSF 676

RESULT 8
 MT30_YEAST STANDARD; PRT; 640 AA.
 ID MT30_YEAST
 AC F39014;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)

15-SEP-2003 (Rel. 42, Last annotation update)
 OT MET30 protein.
 DE MET30 OR Y11046W.
 SN Saccharomyces cerevisiae (Baker's yeast).
 SC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 CX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=X2180-1A;
 RX MEDLINE=36089360; PubMed=8524217;
 RA Thomas D., Kuras L., Barbey R., Cherest H., Blaiseau P.L.,
 RA Surdin-Kerjan Y.
 RT "Met30p, a yeast transcriptional inhibitor that responds to S-
 RT adenosylmethionine, is an essential protein with WD40 repeats."
 RL Mol. Cell. Biol. 15:6526-6534(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX PubMed=9169870;
 RA Churcher C.M., Bowman S., Badcock K., Bankier A., Brown D.,
 RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin M.,
 RA Harris D.E., Horsnell T., Hunt S., Jagels K., Jones M., Lye G.,
 RA Meule S., Odell C., Pearson D., Rejandream M.A., Rice P., Rowley N.,
 RA Skelton J., Smith V., Walsh S., Whitehead S., Barrrell B.G.;
 RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX."
 RL Nature 387:84-87(1997).
 CC -1- FUNCTION: NEGATIVELY REGULATES SULFUR AMINO ACIDS BIOSYNTHESIS
 CC GENES EXPRESSION.
 CC -1- SUBUNIT: SEEMS TO INTERACT WITH MET4.
 CC -1- SIMILARITY: Contains 1 F-box domain.
 CC -1- SIMILARITY: Contains 8 WD repeats.
 CC -1- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT
 CC PROTEINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; Z46861; CAA86905.1; -;
 CC EMBL; L26505; AAA96717.1; -;
 CC PIR; S49932; S49932.
 CC SGD; S0001308; MET30.
 CC InterPro; IPR001810; F-box.
 CC InterPro; IPR001680; WD40.
 CC Pfam; PF00646; F-box; 1.
 CC Pfam; PF00400; WD40; 7.
 CC PRINTS; PR00320; GPROTEINBRPT.
 CC ProDom; PD000018; WD40; 1.
 CC SMART; SM00256; FBOX; 1.
 CC SMART; SM00320; WD40; 6.
 CC PROSITE; PS50181; FBOX; 1.
 CC PROSITE; PS00678; WD_REPEATS_1; 4.
 CC PROSITE; PS00682; WD_REPEATS_2; 6.
 CC PROSITE; PS50294; WD_REPEATS_REGION; 1.
 CC Transcription regulation; Methionine biosynthesis;
 CC Cysteine biosynthesis; Repeat; WD repeat.
 CC DOMAIN 181 227 F-BOX.
 CC REPEAT 300 328 WD 1.
 CC REPEAT 340 368 WD 2.
 CC REPEAT 380 408 WD 3.
 CC REPEAT 419 449 WD 4.
 CC REPEAT 461 499 WD 5.
 CC REPEAT 509 538 WD 6.
 CC REPEAT 550 578 WD 7.
 CC REPEAT 607 635 WD 8 (POTENTIAL).
 CC CONFLICT 61 61 M -> I (IN REF. 1).
 CC SEQUENCE 640 AA; 72835 MW; 5135D4BCA2E1B97 CRC64;
 SQ

Query Match 18.0%; Score 545; DB 1; Length 640;
 Best Local Similarity 28.9%; Pred. No. 1.3e-33;
 Matches 167; Conservative 85; Mismatches 202; Indels 124; Gaps 17;
 QY 34 IPEKNSLRQYNSCARCLN-QETVCLASTAMKTEVCVAK-TKLANGTSSMIVPKORKLS 91
 DB 86 LPEYNFTKCYRHNPDIOFSPHTACYPKQDLKRTQINANIAPLQEQSDIHHIISKYS 145
 QY 92 ASYEKEKELCVKYPEOWSESQVEFVEHLISQCHYGHINSYKPMLOREITAPAR 151
 DB 146 NSNDKIRKL-----LDGILSTSCFPQLSYISLVTHMKIDFISLTP-- 188
 QY 152 GLDIAENILSYLDAKSLCAELVCKEYRVTSDGLMKKLIEMVVRTDSLRLGLAERR- 210
 DB 189 --QELSLKILSYLDCQLCNATRCRKWKQLADDDRVVHMCBOH-----DRKC 236
 QY 211 ---GWGQYLFKNK-----PPDGNAPNSFYRALYPKLIQDIETIESNNRCGHSILQ 258
 DB 237 PNCGWGUPLLHMKRARIQNSTGSSNADIQTTPRPWKVIYRERPKVESNRKNG----- 291
 QY 259 RIHCRSETSK----GVYCLQYDQKIVSGLRDNTIKIDKNTILECKRIILTGTGSLCLQ 314
 DB 292 --HCRIQEFKGMGVLTLOFNVELLFTGSDYSTIGWDLFTGKLIIRLSGSHSDGVKTYL 349
 QY 315 YDERVLIITGSSDSIVRVWDVNTGEMLNTLIHCEAVLHLPFNNGMMVTCKDSRTAVWDM 374
 DB 350 FDDRKLITGSLDKTIRVMNYITGECISTYRGHSDSVLSVDYQKIVSGSADTKVKVHEV 409
 QY 375 ASPDTIDILRRVLVGHRAAVNVDFDQKI--VSGDRTIKVWNTSTCFVRTLNGH--- 429
 DB 410 ESRCTYLR---GHEWVNCVKLHPKSPSCSDDTIRWMDITNSCLKVFRGHVQ 465
 QY 430 KRGIAQCYLRD--RLVSSGSS----- 448
 DB 466 VKIIPITKIDVENLATDNTSDGSSPQDDPTMTDGADESPTPSNEQTVLDENIPYPTHL 525
 QY 449 -----DNTIRLWDECCACLRVLEGHEELVRCIRFONKRVISGAYGKIKWDLVAALDP 503
 DB 526 LSCGLDNTIKLWDVTKCKIRCTQFGHVEGWDAADNFRIISGSHSGSIKWVDLQSG--- 582
 QY 504 RAPAGTLCRLTVHSHGSRVFLQDFEQIVSSSHDDTI 541
 DB 583 -----KCMHTF---NGR--RLQRETOHTQTQSLGDKV 609

RESULT 9
 SC02_NEUCR
 ID SC02_NEUCR STANDARD; PRT: 650 AA.
 AC Q01277;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Sulfur controller-2 (SC02).
 GN SC0N-2.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 CC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=74-OR23-1A / FGSC 987;
 RX MEDLINE=95241499; PubMed=7724564;
 RA Kumar A., Paletta J.V.;
 RT "The sulfur controller-2 negative regulatory gene of Neurospora
 RT crassa encodes a protein with beta-transducin repeats."
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3343-3347(1995).
 CC -1- FUNCTION: NEGATIVELY REGULATES SULFUR STRUCTURAL GENE EXPRESSION.
 CC -1- INDUCTION: EXPRESSED ONLY UNDER LOW-SULFUR CONDITIONS.
 CC -1- SIMILARITY: Contains 8 WD repeats.
 CC -1- SIMILARITY: Contains 1 F-box domain.
 CC -1- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT
 CC PROTEINS.


```

RX MEDLINE=96009891; PubMed=7557402;
RA Saue S., Turcq B., Begueret J.;
RT "A gene responsible for vegetative incompatibility in the fungus
RI Podospira anserina encodes a protein with a GTP-binding motif and G
RL beta homologous domain."
CC -!- FUNCTION: RESPONSIBLE FOR VEGETATIVE INCOMPATIBILITY THROUGH
CC SPECIFIC INTERACTIONS WITH DIFFERENT ALLELES OF THE UNLINKED GENE,
CC HET-C.
CC -!- SIMILARITY: Contains 10 WD repeats.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L28125; AAA85775.1; .
CC PIR; T18521; T18521.
CC InterPro; IPR007111; NACHT_NTPase.
CC InterPro; IPR001680; WD40.
CC Pfam; PF00400; WD40; 10.
CC PRINTS; PR00320; GPROTEINRPT.
CC ProDom; PD000018; WD40; 10.
CC SMART; SM00320; WD40; 10.
CC PROSITE; PS00837; NACHT; 1.
CC PROSITE; PS00678; WD REPEATS 1; 10.
CC PROSITE; PS00082; WD REPEATS 2; 10.
CC PROSITE; PS02094; WD REPEATS_REGION; 1.
CC GTP-binding; Repeat_WD repeat.
CC
CC DOMAIN 294 629 NACHT.
FT NP_BIND 300 307 GTP (POTENTIAL).
FT REPEAT 839 869 WD 1.
FT REPEAT 881 911 WD 2.
FT REPEAT 923 953 WD 3.
FT REPEAT 965 995 WD 4.
FT REPEAT 1007 1037 WD 5.
FT REPEAT 1049 1079 WD 6.
FT REPEAT 1091 1121 WD 7.
FT REPEAT 1133 1163 WD 8.
FT REPEAT 1175 1205 WD 9.
FT REPEAT 1217 1247 WD 10.
SQ SEQUENCE 1356 AA; 149765 MW; 965FB319844E0651 CRC64;

Query Match 15.0%; Score 455.5; DB 1; Length 1356;
Best Local Similarity 31.7%; Pred. No. 26-26;
Matches 121; Conservative 64; Mismatches 131; Indels 66; Gaps 15;

QY 226 APPNSFYRLA-----PKIIQDIETIESNRCGRHSL-----QRIHCRS- 264
DB 802 APTDSMTIKFKXEPGNWISTISVVEAEWNACTQTLGHSVLSVAFSADGQVAFSGSD 861
QY 265 -----ETSG-----VYCLOY--DDQKIVSLGRDNTIKWTKLTKRIL 303
DB 862 DKTIKWDTAGTQTQTLGHSVWVAFSPDRERVASGSDDKTIKIWDAAAGTCTQTL 921
QY 304 TGTGTVSLCLOY--DERVIITGSDSTVRVVDVNTGEMTLIHCEAVLHFRFN-NGNM 360
DB 922 EGHGRGVSVAFSPDQGVASGSDDKTIKIWDAAAGTCTQTLGHSVLSVAFSPDQGR 981
QY 361 VTC-SKDSIAVWNASPTDITLRVLVGHRAAVVVDFF--DKYIVVAGSDRTIKWNT 417
DB 982 VASGSDGDKTIKWDTAGS-----TCTQTLGHSVWVAFSPDQGVASGSDDKTIKIWD 1038
QY 418 STCEPVRTLNHGKRGIACLQYR--DRLVWSSSDNTIRLWDIECGACLRVLGHEELVRC 475
DB 1039 ASGCTQTLGHSVWVAFSPDQGVASGSDDKTIKIWDVAGTCTQTLGHSVWV 1098
QY 476 IRF--DNKRIVSGYDGIKIVMDLVAALDPAPAGTLCRLTLVHSGRVRLOF--DEFO 531
```

```

DB 1099 VAFSPDQGVASGSDGTIKIWD-----AASGT-CTQTLGHSVWVAFSPDQGR 1149
QY 532 IVSSSHDITLWDFLNDPAQ 553
DB 1150 VASGSDGTIKIWDAAAGTCTQ 1171

RESULT 13
POPI SCHPO
ID _POPI_SCHPO STANDARD; PRT; 775 AA.
AC P37060; Q9P7P3;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE WD-repeat protein popl.
GN POP1 OR SPBC1718.01 OR SPBC2G2.18.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=97347242; PubMed=9203581;
RA Kominami K., Toda T.;
RT "Fission yeast WD-repeat protein popl regulates genome ploidy through
RT ubiquitin-proteasome-mediated degradation of the CDK inhibitor Rumi
RT and the S-phase initiator Cdc18."
RL Genes Dev. 11:1548-1560(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Woltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Eger P., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Snpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
RN [3]
RP SUBUNIT.
RC STRAIN=972;
RX MEDLINE=99144318; PubMed=9990507;
RA Kominami K.-I., Ochofrena I., Toda T.;
RT "Two F-box/WD-repeat proteins Pop1 and Pop2 form hetero- and homo-
RT complexes together with cullin-1 in fission yeast SCF (Skip-
RT cullin-1-F-box) ubiquitin ligase."
RL Genes Cells 3:721-735(1998).
RN [4]
RP FUNCTION, SUBUNIT, AND SUBCELLULAR LOCATION.
RX MEDLINE=22325332; PubMed=12167173;
```


use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; X05625; CAA29113.1; -
 CC EMBL; D50617; BAA09229.1; -
 CC EMBL; Z46255; CAA86341.1; -
 CC PIR; S56245; S56245.
 CC PDB; INEX; 18-FEB-03.
 CC SGD; S000185; CDC4.
 CC GO; GO:000082; P-G1/S transition of mitotic cell cycle; IMP.
 CC GO; GO:000086; P-G2/M transition of mitotic cell cycle; IGI.
 CC InterPro; IPR001810; F-box.
 CC Pfam; PF00646; F-box; 1.
 CC Pfam; PF0400; WD40; 6.
 CC PRINTS; PR00320; GPROTEINBRPT.
 CC ProDom; PD000018; WD40; 1.
 CC SMART; SM00256; FBOX; 1.
 CC SMART; SM00320; WD40; 7.
 CC PROSITE; PS50181; FBOX; 1.
 CC PROSITE; PS00678; WD REPEATS; 1; 4.
 CC PROSITE; PS50082; WD REPEATS; 2; 5.
 CC PROSITE; PS50294; WD REPEATS REGION; 1.
 CC Cell division; Mitosis; Sporulation; Repeat; WD repeat; 3D-structure.
 FT DOMAIN 272 319
 FT REPEAT 380 408
 FT REPEAT 420 449
 FT REPEAT 461 493
 FT REPEAT 528 556
 FT REPEAT 568 598
 FT REPEAT 630 658
 FT REPEAT 669 698
 FT REPEAT 779 809
 FT CONFLICT 460 460 K -> E (IN REF. 1).
 FT SEQUENCE 779 AA; 86089 MW; 0348F2F8FA78F3BC CRC64;
 Query Match 13.2%; Score 399; DB 1; Length 779;
 Best Local Similarity 24.3%; Pred. No. 1.8e-22;
 Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;
 QY 56 TVCLASTAMKTENCVAKTKLANGTSSMIVPKQKLSASYEK------ELCVKYPQW 108
 DB 194 TPLAKTKTINN-----NNNIADLESKDSIIISPEVLSDEISAINNLPHAYFK-- 244
 QY 109 SESQVEFEVHLISQCHYGHGINSYKPLORDFTALPARGLDHIAENILSVLDKAS 168
 DB 245 -----NLLFLVANMORSELSDLGLIKDKRLDITLSTPP-----EISLKFNYLQPED 294
 QY 169 LCAAEVCKEYR-VTSDGMLWKKLI--ERMVRTDSLWGLAERRGQGYLFKNKPPDGN 225
 DB 295 IINSLGVSNQNKTIKRSKSLWKKLIISNFV-----SPKGF 331
 QY 226 APPNSFYBALYFKLIQD-----IET--TESNWRGRLSHRIHCRSETSKGVYCLQYD 276
 DB 332 NSLNKLSQKYPKLSQQDRRLSFLNFIKRWYKFPVPTTLRGHMTSVITCLQFE 391
 QY 277 DQKIVSGLRDNTIKWDRNTLECKRILTGTGSLVCLQYD-RVITGSSDSSTVAVWDVN 335
 DB 392 DNYVITGADRRMIRVYDSINKKFLQLSGHDGGVWALKYAHGGLVSGSDTRTVAVWDIK 451
 QY 336 TGEMLNTLHCEAVLHLRFNNGMWTCSDKRSIAVNDMASTDITLRVLVGHAAVNV 395
 DB 452 KG-----CCT-----HVFKGHNSTVRC 468
 QY 396 VDFDD-----KYIVSASGDRITKVMNT-----STCE-----FVRLTNG 428
 DB 469 LDIVEYKNIKIVTGRSDNTLHWKLPKSSVPDGHCEHDYPLVPHFTPEENPYFVGVLRG 528
 QY 429 HKRGIACIQYEDRVVSGSDNTIRWDIECGACIYVLEGHEELVRCIFDN--KGIYGS 486
 DB 529 HMASVTVSGHNVVSGSDNTLIVWDVAQMCKVLISGHTDRIVSTIYDHERKRCISA 588

QY 487 AYDGKIKVWD-----VAALDPRAPAGTL--CLRTLVHSGRVFPLQDFEQIVSS 535
 DB 589 SMDTIRINDLENWNGESYATNSGPAKILGAMYTLQGHIALVGLRLSRKFLVSA 648
 QY 536 SHDITLWDFLNDPAAQAEPSPRSPRTYTY 566
 DB 649 AADGSIRGWD-AND-----YSRKFSY 668
 RESULT 15
 POP2 SCHPO
 ID POP2 SCHPO STANDARD; PRT; 703 AA.
 AC 014170;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 25-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE WD-repeat protein pop2 (Proteolysis factor sud1).
 GN POP2 OR SUD1 OR SPAC4D7.03.
 OS Schizosaccharomyces pombe (Pission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=99227353; PubMed=10209119;
 RA Wolf D.A., McKeon F., Jackson P.K.;
 RT "F-box/WD-repeat proteins pop1p and Sud1p/Pop2p form complexes that
 RL bind and direct the proteolysis of cdc18p.";
 RN Curr. Biol. 9:373-376(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98318628; PubMed=9653157;
 RA Jallepalli P.V., Tien D., Kelly T.J.;
 RT "sud1+ targets cyclin-dependent kinase-phosphorylated Cdc18 and Rum1
 RT proteins for degradation and stops unwanted diploidization in fission
 RL yeast.";
 RN Proc. Natl. Acad. Sci. U.S.A. 95:8159-8164(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=2184401; PubMed=11659360;
 RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sprouks J., Feat N., Hayles S., Baker S., Basham D., Bowman S.,
 RA Collins M., Brown D., Brown S., Chillingworth I., Churcher C.M.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinovitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren I., Whitehead S.,
 RA Woodward J., Voickart G., Aert R., Robben J., Grynolprez B.,
 RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehtach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Gadiu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.I., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Bartell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 RN [4]
 RP FUNCTION, AND SUBUNIT.
 RC STRAIN=972;
 RX MEDLINE=99144318; PubMed=9990507;

RA Kominami K.-I., Ochotorena I., Toda T.;
 RT "Two F-box/WD-repeat proteins Pop1 and Pop2 form hetero- and homo-
 RT complexes together with cullin-1 in fission yeast SCF (Skip-
 RT cullin-1-P-box) ubiquitin ligase.";
 RL Genes Cells 3:721-735(1998).
 RN [5]
 RP FUNCTION, SUBUNIT, AND SUBCELLULAR LOCATION.
 RX MEDLINE=22325332; PubMed=12167173;
 RA Seibert V., Prohl C., Schoutz I., Rhee E., Lopez R., Abderazzaq K.,
 RA Zhou C., Wolf D.A.;
 RT "Combinatorial diversity of fission yeast SCF ubiquitin ligases by
 RT homo- and heterooligomeric assemblies of the F-box proteins Pop1 and
 RT Pop2p.";
 RL BMC Biochem. 3:22-22(2002).
 CC -!- FUNCTION: Involved in maintenance of ploidy through proteasome
 CC dependent degradation of CDK inhibitor rum1 and S-phase initiator
 CC cdc18. Functions as a recognition factor for rum1 and cdc18, which
 CC are subsequently ubiquitinated and targeted to the 26S proteasome
 CC for degradation.
 CC -!- SUBUNIT: Homodimer and heterodimer with pop1. Binds to pcul, pip1
 CC and cdc18.
 CC -!- SUBCELLULAR LOCATION: Nuclear, and cytoplasmic.
 CC -!- SIMILARITY: Contains 1 F-box domain.
 CC -!- SIMILARITY: Contains 6 WD repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR ENBL; AF038867; AAB95480.1; -;
 DR ENBL; AF064515; AAC39496.1; -;
 DR ENBL; Z98602; CAB11275.1; -;
 DR PIR; T43557; T43557.
 DR GeneDB SPombe; SPAC4D7.03; -;
 DR InterPro; IPR001810; F-box.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00646; F-box; 1.
 DR Pfam; PF00400; WD40; 6.
 DR PRINTS; PR00320; GPROTEINERPT.
 DR PRODOM; PD000018; WD40; 1.
 DR SMART; SM00256; FBOX; 1.
 DR SMART; SM00320; WD40; 7.
 DR PROSITE; PS50181; FBOX; 1.
 DR PROSITE; PS00678; WD_REPEATS_1; 3.
 DR PROSITE; PS50082; WD_REPEATS_2; 6.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Nuclear protein; Repeat; WD repeat.
 FT DOMAIN 236 283
 FT REPEAT 389 417
 FT REPEAT 429 473
 FT REPEAT 505 533
 FT REPEAT 545 575
 FT REPEAT 587 615
 FT REPEAT 625 654
 FT REPEAT 703 736
 SQ SEQUENCE 703 AA; 79605 MW; 0336A0568C152396 CRC64;
 Query Match 13.1%; Score 396.5; DB 1; Length 703;
 Best Local Similarity 23.3%; Pred. NO. 2.4e-22;
 Matches 146; Conservative 92; Mismatches 226; Indels 163; Gaps 25;
 QY 14 KFMNRRERD-CNNGEPPEPKI-----IPEKNSLRQYNSCARLCL----NQETVCL 59
 DB 96 KFNESIVSVKLDHDPDPSPGDFSIISLVPOKNPL-YSHSSLPKPKIISIDRNRIKLD 154
 QY 60 ASTAMKTENCVAKTKLANGTSSMIVPKORKLSAYEK-EKELCVKYFQWSESQVEFVE 118
 DB 155 NSISSNDNFPSPKPV--DTSNTVSPGSPISLEDLEDNLQSIQVTFEDLPFGIQ-SYAP 211
 QY 119 HLIQCMCHYQGHG-INSYLKPMQLQDFITALPARGLDHAENILSYLDAKSLCAELVCK 177

DB 212 FQLLRSCNRMQLLNECEPLKKDILNSLPF-----SIVOSILLNLDIHSLSCRLVSP 267
 QY 178 EWRVY-----TSDGMLWKKLIERWVRTDSLWRGLABRRGQGYLFKNKPPDGNAPP---N 229
 DB 268 TWRILDVHTS---YWKHMF-----SLFGQIENEDW-----KYANFNLRPPFLHND 312
 QY 230 SFYZALVPKIIQDIETIESNWRGHSIOR-----
 DB 313 QISDDYPEPEIFK-----RHFLNRKRWLPSPPPSHLSFPPIHVPNFMITSLLH 360
 QY 260 -----IHKSETSK-GVYCLQYDQKIVSGLRDNTIKIMDKNT 296
 DB 361 KDRITTSGGTIIQHNATIGVLEARLEHKGWAVKIHENTLVSGSIDKTVRVWNIEK 420
 QY 297 LECKRILTGTGVLQLOY-----DERVLIITGSSDSTVRYVDYNTGEML 340
 DB 421 AKCTHIFRGHISIIIRCEILVPSRLIRHGVEIVEPDQPIYVSGSRDHTLRVW----- 472
 QY 341 NTLIHCEAVLHFRFNNGMVTCSDRSIAVWDMASPTDITLRLRVLGHRAAVNVDFDD 400
 DB 473 -----KLPKNTDPPYLPDNTNSIDRWE-KNPYFV---HTLIGHTDSVRTISYG 517
 QY 401 KYIVSASGDETIKVMNTSTCEFVRLNGHKRGCIACLOYRD--RLVVGSSDNTIELWDIE 458
 DB 518 DILVSGSYDSIRKINRVSTGECYHURGHSIRIYSLVPEPERNICISGMDKSIKRVWDL 577
 QY 459 CGACLRVLEGHEELVRCIRFDNKNRIVSGAYDGKIKVWDLVAALDPRAPAGTLCRLTIVEH 518
 DB 578 TGTCKYVLEGHDAPVTLLNVFQNRLLISGSADSTIRWD---LNTGKP-----LMVLPSN 628
 QY 519 SGRVRLQDFEFQIVSSSHDDTLIWD 545
 DB 629 SGYISSFVSDEHKII-SGNDGSKVCLWD 654

Search completed: October 22, 2003, 10:23:22
 Job time : 32 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

1M protein - protein search, using sw model

Run on: October 22, 2003, 10:08:36 ; Search time 84 seconds

(without alignments)
1075.184 Million cell updates/sec

Title: US-09-601-168B-2

Perfect score: 3034
Sequence: 1 MDPABAVLQKALKFMNSE.....PAAQAEPPSPSTTYTISR 569

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_13Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3034	100.0	569	20 AAY24054	A human beta-trans
2	3034	100.0	569	21 AAB12813	Human beta-transdu
3	3034	100.0	569	21 AAY96697	Human beta-TrCP.
4	3034	100.0	569	21 AAY83041	F-box protein FBP-
5	3034	100.0	569	21 AAY83250	F-box protein hbet
6	3034	100.0	569	21 AAY44249	Human cell signal
7	3034	100.0	569	23 AAO22446	Human F-box protein
8	3034	100.0	569	23 ABG69473	Human bait protein
9	3027	99.8	569	22 AAB48298	Human Fli1 protein

10	3027	99.8	608	22	AAW00960	Human bone marrow
11	3006	99.1	605	22	AAW78582	Human protein SEQ
12	2997	98.8	569	21	AAW12812	Mouse ubiquitin li
13	2997	98.8	569	21	AAW83254	F-box protein FWDI
14	2992.5	98.6	632	22	AAW78584	Human protein SEQ
15	2989.5	98.5	654	22	AAW79566	Human protein SEQ
16	2989.5	98.5	654	22	AAW79567	Human protein SEQ
17	2989.5	98.5	654	22	AAW79568	Human protein SEQ
18	2962	97.6	590	22	AAW00847	Human bone marrow
19	2882	95.0	579	22	AAW78583	Human protein SEQ
20	2882.5	95.1	517	16	AAW85852	WD-40 domain-contg
21	2884.5	78.6	542	21	AAW96696	Human E3 ubiquitin
22	2884.5	78.6	542	22	AAW79127	Human protein SEQ
23	2884.5	78.6	542	22	AAW40208	Human polypeptide
24	2884.5	78.6	550	22	AAW41994	Human polypeptide
25	2089.5	68.9	510	22	ABW59857	Drosophila melanog
26	1163	38.3	219	23	AAW98087	Drosophila melanog
27	989.5	32.6	265	22	AAU86942	Human beta TrCP (b
28	935	30.8	448	22	AAW88289	Human DNA repair a
29	661	21.8	1326	22	ABW67237	Beta-TrCP.N/SKP2.C
30	661	21.8	1326	22	ABW67238	Drosophila melanog
31	661	21.8	1326	22	ABW70051	Drosophila melanog
32	644.5	21.2	626	20	AAW22469	Human smyc-N-sei-1
33	644.5	21.2	626	22	AAW59201	Protein encoded by
34	640	21.1	540	20	AAW22465	Human hippocampal
35	640	21.1	540	20	AAW22466	Human mammary sei-
36	640	21.1	540	20	AAW59197	Human hippocampal
37	640	21.1	540	22	AAW59200	Human mammary sei-
38	640	21.1	545	20	AAW22464	Human hippocampal
39	640	21.1	545	22	AAW59196	Human hippocampal
40	640	21.1	553	20	AAW22463	Human hippocampal
41	640	21.1	553	22	AAW93475	Human protein sequ
42	640	21.1	553	22	AAW59195	Human hippocampal
43	640	21.1	559	20	AAW22467	Human mammary sei-
44	640	21.1	559	22	AAW59199	Human mammary sei-
45	640	21.1	589	20	AAW22466	Human mammary sei-

ALIGNMENTS

RESULT 1
AAY24054
ID AAY24054 standard; Protein; 569 AA.

AC AAY24054;
XX
XX
DT 20-MAR-2003 (updated)
DT 30-SEP-1999 (first entry)
XX
XX
DE A human beta-transducin repeat containing protein.

KW Beta-transducin repeat containing protein; beta-TrCP; Skp1p;
KW proteasome degradation pathway; Vpu protein; beta-catenin;
KW human immune deficiency virus-1; HIV-1; cellular protein; IkappaB;
KW ubiquitination; phosphorylated protein; tumour; apoptosis; Alzheimer's;
KW antiviral; antitumour; cell cycle regulation; protein degradation;
KW and anti-inflammatory; osteo-articular inflammation; acute inflammation;
KW tumour necrosis factor.

OS Homo sapiens.

XX
FH Key Location/Qualifiers
FT Region 147..191
FT /note= "F box sequence"
FT Region 259..292
FT /note= "WD motif"
FT Region 304..332
FT /note= "WD motif"
FT Region 343..372
FT /note= "WD motif"
FT Region 387..415
FT /note= "WD motif"

Region 427..455
/note= "WD motif"
Region 467..492
/note= "WD motif"
Region 516..544
/note= "WD motif"

40938969-A1.

05-AUG-1999.

29-JAN-1999; 99WO-PR00196.

29-DEC-1998; 98FR-0015545.
0-JAN-1998; 98FR-0001100.

INSP } INST PASTEUR.
INRM } INST NAT SANTE & RECH MEDICALE.

renzana Seidedos F, Benarous R, Concordet J, Durand H;
Roli M, Margottin F;

PI; 1999-469329/39.

-PSDB; AAX86501.

ew human beta-transducin repeat containing protein and its
ragments useful as, or to screen for, antiviral, antitumour,
nti-inflammatory and anti-Alzheimer's agents

laim 1; Page 60-61; 71pp; French.

he present sequence represents a human beta-transducin repeat containing
rotein (beta-TrCP). The protein directs proteins to the proteosome
gradation pathways. The protein is able to interact with the Vpu
rotein of human immune deficiency virus-1 (HIV-1), cellular proteins
appab or beta-catenin (bc) and/or protein Skp1. The protein controls
liquinylation of phosphorylated proteins and thus their targeting to
osomes for degradation. Depending on whether the process is
inhibited or promoted, the result may be delayed breakdown of Cb4 (in
ses of HIV-1 infection); increased activity of Ikb (and thus reduced
tivity of NFkappaB) and increased degradation of mutant bc in tumour
ils, or increased bc survival (and reduced apoptosis) in Alzheimer's
tients. The beta-TrCP protein, and its active peptide fragments, or its
cleic acid, are used to screen for anti HIV-1 agents (antivirals),
t tumour agents that disrupt cell cycle regulation or protein
gradation in human tumour cells, and anti-inflammatory agents that
rupt activation by NFkappaB. Fragments of the protein are also
etral for treating osteo-articular inflammation or acute inflammation
sociated with release of tumour necrosis factor.
dated on 20-MAR-2003 to correct PA field.)

quence 569 AA;

Match 100.0%; Score 3034; DB 20; Length 569;
Local Similarity 100.0%; Pred. No. 2.5e-286;
es 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MDPAEAVLQEKALFPMNSEREDCNQGEPRKIIPEKNSLRQTYNSCARLCLNQETVCLA 60
1 MDPAEAVLQEKALFPMNSEREDCNQGEPRKIIPEKNSLRQTYNSCARLCLNQETVCLA 60

61 STAMKTEVCVAKTLANGTSSMIVPKOKLSASVEKEKELCVKYPFOWSESQVFEVHL 120

61 STAMKTEVCVAKTLANGTSSMIVPKOKLSASVEKEKELCVKYPFOWSESQVFEVHL 120

121 ISQMCHYQGHINSYKPMLODFITAPARGLDHIAENTLSYLDAKSLCAELVCKSWY 180

121 ISQMCHYQGHINSYKPMLODFITAPARGLDHIAENTLSYLDAKSLCAELVCKSWY 180

181 RVTSDGLWKKLIERMVTDSLWGLAERRGQGLPKPKPPDGNAPPNSFYRALYPKII 240

181 RVTSDGLWKKLIERMVTDSLWGLAERRGQGLPKPKPPDGNAPPNSFYRALYPKII 240

QY 241 QDIETIESNWRGCRHSLQRIHCFSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300
DB 241 QDIETIESNWRGCRHSLQRIHCFSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300
QY 301 RILTGHGTSVLCLOYDERVITIGSSDSSTVYVMDVNTGEMWNTLIHHCEAVLHLRFNNGMM 360
DB 301 RILTGHGTSVLCLOYDERVITIGSSDSSTVYVMDVNTGEMWNTLIHHCEAVLHLRFNNGMM 360
QY 361 VTCSDKRSIAVMDVASPTDITLRRVLVGHRAAVNVVDFDDKYIVSASGDRTIKVMNTSTC 420
DB 361 VTCSDKRSIAVMDVASPTDITLRRVLVGHRAAVNVVDFDDKYIVSASGDRTIKVMNTSTC 420
QY 421 EFVETLNGHKRGIACTQYRDRLVYVSGSSDNTTIRLWDIECGACLRVLEGHEELVRCIRFDN 480
DB 421 EFVETLNGHKRGIACTQYRDRLVYVSGSSDNTTIRLWDIECGACLRVLEGHEELVRCIRFDN 480
QY 481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCLRTLVEHSGRVFRLOQDFEQIVSSSHDDT 540
DB 481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCLRTLVEHSGRVFRLOQDFEQIVSSSHDDT 540
QY 541 ILIWDFLNDPAAQAEPRSPSRITYTISR 569
DB 541 ILIWDFLNDPAAQAEPRSPSRITYTISR 569

RESULT 2

AAB12813

ID AAB12813 standard; protein; 569 AA.

AC AAB12813;

XX 27-NOV-2000 (first entry)

DE Human beta-transducin repeat containing protein (beta-TrCP) SEQ ID NO:3.

XX Ubiquitin ligase SCF complex; F-box protein; ubiquitination; IkappaB;

KW beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FMD1;

KW gene therapy; colon cancer; beta-transducin repeat containing protein;

KW beta-TrCP.

XX Homo sapiens.

XX JP2000166542-A.

XX 20-JUN-2000.

XX 02-DEC-1998; 98JP-0343437.

XX 02-DEC-1998; 98JP-0343437.

XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

XX WPI; 2000-485550/43.

XX N-PSDB; AAA73132.

XX F-box protein of ubiquitin ligase SCF complex which promotes the

XX ubiquitination of IkappaB or beta-catenin

XX Claim 3; Page 10-12; 19pp; Japanese.

XX The present invention describes an F-box motif protein of ubiquitin
CC ligase SCF complex which promotes the ubiquitination of IkappaB or
CC beta-catenin and is constituted by Skp1 protein, Cull1 protein and a
CC complex (SCF complex) of F-box protein containing F-box motif and WD40
CC repeat motif and has the amino acid sequence of 45 residues (AAB12811)
CC or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin
CC ligase FMD1 protein) and (AAB12813, which is human beta-transducin
CC repeat containing protein (beta-TrCP)). The F-box protein can be used for
CC the gene therapy of colon cancer by being recombinated to a virus vector.

XX Sequence 569 AA;

Query Match 100.0%; Score 3034; DB 21; Length 569;

Local Similarity 100.0%; Pred. No. 2.5e-286;
Res 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MDPAAVLQEKALFNMSSEREDCNNGEPKRIIPEKNSLRQTYNSCARLCLNQETVCLA 60
1 MDPAAVLQEKALFNMSSEREDCNNGEPKRIIPEKNSLRQTYNSCARLCLNQETVCLA 60
61 STAMKTENCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKYPQWSESQVFEVHL 120
61 STAMKTENCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKYPQWSESQVFEVHL 120
121 ISOMCHYQHGHINSYLPKMLQROFITALPARGLDHIAENILSYLDAKSLCAELVCKEWY 180
121 ISOMCHYQHGHINSYLPKMLQROFITALPARGLDHIAENILSYLDAKSLCAELVCKEWY 180
181 RVTS DGMWLKXLIERMVTDLSMRGLAERGGQYLFKPKPPDGNAPPNSFYRALYPKII 240
181 RVTS DGMWLKXLIERMVTDLSMRGLAERGGQYLFKPKPPDGNAPPNSFYRALYPKII 240
241 QDIETIESNRWRCGRHSLSQRIHCRSETSKGVYCLQYDDQKIVSGLRDNITIKIWDKNTLECK 300
241 QDIETIESNRWRCGRHSLSQRIHCRSETSKGVYCLQYDDQKIVSGLRDNITIKIWDKNTLECK 300
301 RILTGHTGSVLCQYDERVITGSSDSTVRVMDVNTGEMTLNLIHCEAVLHLRFNNGMM 360
301 RILTGHTGSVLCQYDERVITGSSDSTVRVMDVNTGEMTLNLIHCEAVLHLRFNNGMM 360
361 VTCSDKRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDKYIVSAGDRTIKVWNTSTC 420
361 VTCSDKRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDKYIVSAGDRTIKVWNTSTC 420
421 EFVRTLNGHKRGIAKLVMDLVAALDPRAPAGTLCRLTVEHSGRVRFLQDFEQIVSSSHDDT 540
421 EFVRTLNGHKRGIAKLVMDLVAALDPRAPAGTLCRLTVEHSGRVRFLQDFEQIVSSSHDDT 540
481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTVEHSGRVRFLQDFEQIVSSSHDDT 540
481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTVEHSGRVRFLQDFEQIVSSSHDDT 540
541 ILIWDFLNDPAAQAEPPRSPRTYTYISR 569
541 ILIWDFLNDPAAQAEPPRSPRTYTYISR 569

96697;
96697 standard; Protein; 569 AA.

-SEP-2000 (first entry)
nan beta-TrCP.
ubiquitin ligase; beta-TrCP; P-box; WD protein; I-kappa-B; inhibitor;
nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;
i-inflammatory; immunosuppressive; cytostatic.
homo sapiens.

00034447-A2.
JUN-2000.1
DEC-1999; 99WO-US29371.
DEC-1998; 98US-0210060.
GN-) SIGNAL PHARM INC.
SS) YISSUM RES & DEV CO.
ning AM, Mercurio F, Amit S, Ben-eriah Y, Davis M, Hatzubai A;
on I, Yaron A;

XX WP; 2000-431294/37.
DR N-ESDB; AAA51229.
XX polypeptide enhancing phosphorylated I-kappaB ubiquitination useful for
PT treating disorder associated with NF-kappaB activation e.g. cancer,
PT comprising amino acid sequence of human E3 ubiquitin ligase or its
PT variant
XX Claim 21; Page 72-74; 77pp; English.
XX Human beta-TrCP, an F-box/WD protein family member, has been shown to
CC have homology to human E3 ubiquitin ligase (E3). E3 enhances
CC ubiquitination of phosphorylated I-kappa-B, an inhibitor protein of
CC nuclear factor kappa-B (NF-kappa-B). Understanding I-kappa-B
CC degradation via the ubiquitin pathway is useful for identifying
CC modulators of NF-kappa-B. In vitro analysis suggests that deletion of
CC activation of NF-kappa-B. In vitro analysis suggests that deletion of
CC the F-box results in a protein that functions as a dominant negative
CC molecule in vivo. Transient over-expression of delta-beta-TrCP (a
CC deletion mutant) inhibited the degradation of endogenous I-kappa-B-alpha
CC in stimulated Jurkat cells, resulting in accumulation of phosphorylated
CC I-kappa-B-alpha. E3 can be used to screen for modulators of NF-kappa-B
CC activity. E3 and beta-TrCP can be used to modulate NF-kappa-B to treat
CC inflammatory diseases, autoimmune diseases, cancer and viral infections.
XX Sequence 569 AA;
SQ

Query Match 100.0%; Score 3034; DB 21; Length 569;
Best Local Similarity 100.0%; Pred. No. 2.5e-286;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAAVLQEKALFNMSSEREDCNNGEPKRIIPEKNSLRQTYNSCARLCLNQETVCLA 60
Db 1 MDPAAVLQEKALFNMSSEREDCNNGEPKRIIPEKNSLRQTYNSCARLCLNQETVCLA 60
QY 61 STAMKTENCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKYPQWSESQVFEVHL 120
Db 61 STAMKTENCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKYPQWSESQVFEVHL 120
QY 121 ISOMCHYQHGHINSYLPKMLQROFITALPARGLDHIAENILSYLDAKSLCAELVCKEWY 180
Db 121 ISOMCHYQHGHINSYLPKMLQROFITALPARGLDHIAENILSYLDAKSLCAELVCKEWY 180
QY 181 RVTS DGMWLKXLIERMVTDLSMRGLAERGGQYLFKPKPPDGNAPPNSFYRALYPKII 240
Db 181 RVTS DGMWLKXLIERMVTDLSMRGLAERGGQYLFKPKPPDGNAPPNSFYRALYPKII 240
QY 241 QDIETIESNRWRCGRHSLSQRIHCRSETSKGVYCLQYDDQKIVSGLRDNITIKIWDKNTLECK 300
Db 241 QDIETIESNRWRCGRHSLSQRIHCRSETSKGVYCLQYDDQKIVSGLRDNITIKIWDKNTLECK 300
QY 301 RILTGHTGSVLCQYDERVITGSSDSTVRVMDVNTGEMTLNLIHCEAVLHLRFNNGMM 360
Db 301 RILTGHTGSVLCQYDERVITGSSDSTVRVMDVNTGEMTLNLIHCEAVLHLRFNNGMM 360
QY 361 VTCSDKRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDKYIVSAGDRTIKVWNTSTC 420
Db 361 VTCSDKRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDKYIVSAGDRTIKVWNTSTC 420
QY 421 EFVRTLNGHKRGIAKLVMDLVAALDPRAPAGTLCRLTVEHSGRVRFLQDFEQIVSSSHDDT 540
Db 421 EFVRTLNGHKRGIAKLVMDLVAALDPRAPAGTLCRLTVEHSGRVRFLQDFEQIVSSSHDDT 540
QY 481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTVEHSGRVRFLQDFEQIVSSSHDDT 540
Db 481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTVEHSGRVRFLQDFEQIVSSSHDDT 540
QY 541 ILIWDFLNDPAAQAEPPRSPRTYTYISR 569
Db 541 ILIWDFLNDPAAQAEPPRSPRTYTYISR 569

4
1
AY83041 standard; Protein: 569 AA.
AY83041;
5-AUG-2000 (first entry)
-box protein FBP-1.
box protein; FBP; diagnosis; treatment; screening; agonist;
agonist; proliferative disorder; differentiative disorder;
-east cancer; prostate cancer; ovarian cancer; cancer;
-all cell lung carcinoma; immune disorder; cardiovascular disorder;
inflammatory disorder; human.
mo sapiens.
200012679-A1.
-MAR-2000.
-AUG-1999; 99WO-US19560.
-AUG-1998; 98US-0098355.
-FEB-1999; 99US-0118568.
-MAR-1999; 99US-0124449.
YNY) UNIV NEW YORK STATE.
aur DS, Pagano M, Latres E;
I; 2000-256635/22.
PSDB; AAZ93350.
vel nucleic acid for screening compounds useful for treating
oliferative and differentiative disorders such as cancer and immune
sorders comprises sequences encoding ubiquitin ligases -
closure; Figure 3a; 245pp; English.
leic acids encoding substrate-targeting subunits of ubiquitin
ases with F-box motifs (F-box proteins) are useful for diagnosis
proliferative and differentiated related disorders by measuring
gene expression. Cells expressing such proteins or
ir fragments are useful for screening compounds. The compounds
agonists or antagonists, which are useful for treating a
liferative or differentiative disorder in a mammal such as
est, ovarian and prostate cancer and small cell lung carcinoma
also major opportunistic infections, immune disorders,
divascular diseases and inflammatory disorders. FBP protein,
logs, derivatives and their subsequences, anti-FBP antibodies
also useful in diagnosis of the disorders.
quence 569 AA;
Match 100.0%; Score 3034; DB 21; Length 569;
ocal Similarity 100.0%; Pred. No. 2.5e-286;
s 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MDPAAVLQKALFPMNSRDCNNGEPKRIIPKNSURQYNSCARLCLNQETVCLA 60
1 MDPAAVLQKALFPMNSRDCNNGEPKRIIPKNSURQYNSCARLCLNQETVCLA 60
61 STAMKTCNVAKTLANGTSSMIVPKOKLSAYEKEKELCVKFEQWSESDQVEFVHL 120
61 STAMKTCNVAKTLANGTSSMIVPKOKLSAYEKEKELCVKFEQWSESDQVEFVHL 120
121 ISQCHYQHGHINSYKPMQLORDITAPARGLDHIAENILSYLDAKSLCAELVCKEWY 180
121 ISQCHYQHGHINSYKPMQLORDITAPARGLDHIAENILSYLDAKSLCAELVCKEWY 180
181 RVTSDGMLWKKLIERMVRTDSLWRGLAERRGWGYLFKNKPPDGNAPPNSFYRALYPKII 240

Db 181 RVTSDGMLWKKLIERMVRTDSLWRGLAERRGWGYLFKNKPPDGNAPPNSFYRALYPKII 240
QY 241 QDIETIESNRCGRHSRLRIHCRSETSGVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300
Db 241 QDIETIESNRCGRHSRLRIHCRSETSGVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300
QY 301 RLITGHTGSLVCLQYDERVITGSSDSTVRVNDVNTGEMTLTHHCEAVLHLRFNNGMY 360
Db 301 RLITGHTGSLVCLQYDERVITGSSDSTVRVNDVNTGEMTLTHHCEAVLHLRFNNGMY 360
QY 361 VTCSKDRSTAVWDMASPTDITLRVLVGHRAAVNVDFDDKYIVSASGDRITIKVWNTSTC 420
Db 361 VTCSKDRSTAVWDMASPTDITLRVLVGHRAAVNVDFDDKYIVSASGDRITIKVWNTSTC 420
QY 421 EFVRTLNGHKGRTIACLOYRDLVVGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN 480
Db 421 EFVRTLNGHKGRTIACLOYRDLVVGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN 480
QY 481 KRIVSGAYDGKIKVWDLVAALDPRAPAGTLCRLTILVEHSGRVFRLQDFEFOIVSSSHDDT 540
Db 481 KRIVSGAYDGKIKVWDLVAALDPRAPAGTLCRLTILVEHSGRVFRLQDFEFOIVSSSHDDT 540
QY 541 ILIWDPLNDPAAQAEPSPRSPTYIISR 569
Db 541 ILIWDPLNDPAAQAEPSPRSPTYIISR 569
RESULT 5
AAY83250
ID AAY83250 standard; Protein: 569 AA.
XX AC AAY83250;
XX DT 16-AUG-2000 (first entry)
XX DE F-box protein hBetaTrCp.
XX KW Ubiquitin ligase; SCF; F-box protein; targeted degradation;
KW destabilisation; proteolysis; drug discovery; gene therapy; cancer;
KW oncoprotein; Huntington's disease; gene knockout; delivery systems;
XX OS Homo sapiens.
XX EN W0200022110-A2.
XX FD 20-APR-2000.
XX PP 08-OCT-1999; 99WO-US23705.
XX PR 09-OCT-1998; 98US-0103787.
XX PA (HARD) HARVARD COLLEGE.
XX PI Zhou P, Howley P;
XX DR WPI; 2000-317970/27.
XX DR N-PSDB; AAZ933710.
XX PT Targeting degradation of polypeptide useful for treating cancer and
PT other proliferative disorders, involves conjugating polypeptide with
PT ubiquitin protein ligase or inhibiting ubiquitination using organic
XX compound
XX Claim 9; Page 171; 185pp; English.
XX The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin
CC ligases) which can be used for the targeted degradation of a target
CC polypeptide in vivo. Targeted degradation is achieved by expressing
CC the ubiquitin ligase in a cell linked to the interaction domain of
CC the target polypeptide and thereby recruiting the target polypeptide
CC to the ubiquitin ligase. Such methods are useful for decreasing or

increasing the level of a target polypeptide and for creating and expressing a destabilized polypeptide which is subjected to SCF mediated proteolysis. Degrading any desired protein in a cell is useful for preventing or treating diseases caused by the presence of normal amount of the specific polypeptides, for drug discovery and for gene therapy. Diseases treated include cancer, by degradation of oncoproteins, Huntington's disease, other proliferative disorders and microbial infections. The method provides a quick and easy alternative to gene knockout technology. The target polypeptide can be degraded at all stages, or a specific stage, of development in the mature animal.

sequence 569 AA;
Match 100.0%; Score 3034; DB 21; Length 569;
Local Similarity 100.0%; Pred. No. 2,5e-286;
ies 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MDPAEVLQEKALFPMNSREDGNGEPKRIPEKNSLRQYNSCARLCLNQETVCLA 60
1 MDPAEVLQEKALFPMNSREDGNGEPKRIPEKNSLRQYNSCARLCLNQETVCLA 60
61 STAMKTCNCVAKTKLANGTSMVLPKOKLSASVEKEKLCVYFQWSESQOEFVEHL 120
61 STAMKTCNCVAKTKLANGTSMVLPKOKLSASVEKEKLCVYFQWSESQOEFVEHL 120
121 ISQCHYQHGHSINLYKPMQLQRFITLARGLDHIAENILSYLDAKSLCAELVCKEY 180
121 ISQCHYQHGHSINLYKPMQLQRFITLARGLDHIAENILSYLDAKSLCAELVCKEY 180
181 RVTSDGMLWKKLIERMVRTSLMRGLAERGWQYLFKNKPPDGNAPPNSFYRALYPKII 240
181 RVTSDGMLWKKLIERMVRTSLMRGLAERGWQYLFKNKPPDGNAPPNSFYRALYPKII 240
241 QDIETISNWRGHSQRHCRSETSGVYCLQYDDQKIVSGLRDNTIKINDKNTLECK 300
241 QDIETISNWRGHSQRHCRSETSGVYCLQYDDQKIVSGLRDNTIKINDKNTLECK 300
301 RILTHGTGSLVLCLOYDERVITGSSDSTVRVNDVTGEMNTLIHCEAVLHLRFNNGMW 360
301 RILTHGTGSLVLCLOYDERVITGSSDSTVRVNDVTGEMNTLIHCEAVLHLRFNNGMW 360
361 VTCSKDRSIAVMDWASPTDITLRVLVGHRAAVNVDFDDKIVSAGDRTIKVWNTSTC 420
361 VTCSKDRSIAVMDWASPTDITLRVLVGHRAAVNVDFDDKIVSAGDRTIKVWNTSTC 420
421 EFVRTLNGHKGAGIACQYRDLRVVSGSSDNTIRLWDIECGACLRVLEGEELVRCIRFDN 480
421 EFVRTLNGHKGAGIACQYRDLRVVSGSSDNTIRLWDIECGACLRVLEGEELVRCIRFDN 480
481 KRIVSGAYDGKIKVNDLVAALDPAPAGTICLRLTVEHSGRVFRLQDFEQIVSSSHDDT 540
481 KRIVSGAYDGKIKVNDLVAALDPAPAGTICLRLTVEHSGRVFRLQDFEQIVSSSHDDT 540
541 ILIWDFLNDPAAQAEPPRSPRTYTIISR 569
541 ILIWDFLNDPAAQAEPPRSPRTYTIISR 569

44249 standard; Protein; 569 AA.

44249;

PEB-2000 (first entry)

an cell signalling protein-12.

1 signalling protein-12; CSIGP-12; cell proliferation;
lamatory disorder; cirrhosis; cancer; hepatitis; AIDS;
eriosclerosis; Addison's disease; multiple sclerosis.

OS Homo sapiens.
XX Key Location/Qualifiers
FH Modified-site 19
FT Modified-site /note= "Potential phosphorylation site"
FT Modified-site 39
FT Modified-site /note= "Potential phosphorylation site"
FT Modified-site 91
FT Modified-site /note= "Potential phosphorylation site"
FT Modified-site 109
FT Modified-site /note= "Potential phosphorylation site"
FT Modified-site 162
FT Modified-site /note= "Potential phosphorylation site"
FT Modified-site 266
FT Modified-site /note= "Potential phosphorylation site"
FT Modified-site 288
FT Modified-site /note= "Potential phosphorylation site"
FT Modified-site 328
FT Modified-site /note= "Potential phosphorylation site"
FT Modified-site 376
FT Modified-site /note= "Potential phosphorylation site"
FT Modified-site 381
FT Modified-site /note= "Potential phosphorylation site"
FT Modified-site 411
FT Modified-site /note= "Potential phosphorylation site"
FT Modified-site 418
FT Modified-site /note= "Potential phosphorylation site"
FT Modified-site 451
FT Modified-site /note= "Potential phosphorylation site"
FT Modified-site 514
FT Modified-site /note= "Potential phosphorylation site"
FT Modified-site 519
FT Modified-site /note= "Potential phosphorylation site"
FT Modified-site 535
FT Modified-site /note= "Potential phosphorylation site"
FT Modified-site 536
FT Modified-site /note= "Potential phosphorylation site"
FT Modified-site 17
FT Modified-site /note= "Potential glycosylation site"
FT Modified-site 77
FT Modified-site /note= "Potential glycosylation site"
FT Modified-site 416
FT Modified-site /note= "Potential glycosylation site"
FT Region 320..334
FT Region /label= Signature_sequence
FT Region 360..374
FT Region /label= Signature_sequence
FT Region 403..417
FT Region /label= Signature_sequence
FT Region 443..457
FT Region /label= Signature_sequence
FT Region 483..497
FT Region /label= Signature_sequence
FT Region 532..546
FT Region /label= Signature_sequence
XX WO9958558-A2.
XX
XX PD 18-NOV-1999.
XX PF 13-MAY-1999; 99WO-US10567.
XX PR 13-MAY-1998; 98US-0085343.
XX PR 26-AUG-1998; 98US-0098010.
XX
XX PA (INCY-) INCYTE PHARM INC.
XX Bandman O, Hillman JL, Lal P, Yue H, Tang YT, Patterson C;
PI Baughn MR, Yang J;
XX
XX WPI; 2000-086432/07.
DR N-ESDB; AAZ29233.
XX

uman cell signaling proteins useful for, e.g. diagnosing cell
proliferative and inflammatory disorders -
laum 1; Page 77-78; 90pp; English.

he present sequence is cell signalling protein-12 (CSIGP-12) encoded
y cDNA obtained from Incyte clone 3239149 of COLAUCT01 library. It is
xpressed in musculoskeletal, gastrointestinal and nervous tissues and is
ound to be homologous to beta-transducin repeats containing
rotein. Fragments of CSIGP encoding nucleic acid can be used as
ybridisation probe for detecting CSIGP related sequences or allelic
ariants. Recombinant CSIGP can be produced in host cells by transforming
hem with genetically engineered vectors. Agonists or antagonists can be
sed in the treatment of cell proliferative and inflammatory disorders
associated with decreased or increased CSIGP expression. CSIGP is used in
he diagnosis, prevention and treatment of cell proliferative disorders
like arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory
isorders like AIDS, Addison's disease, multiple sclerosis, etc.

sequence 569 AA;

f Match 100.0%; Score 3034; DB 21; Length 569;
Local Similarity 100.0%; Pred. No. 2.5e-286;
es 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MDPAEAVLOEKALFKMNSEREDCNNGEPKIIPEKNSLRQTYNSCARLCLNQETVCLA 60

1 MDPAEAVLOEKALFKMNSEREDCNNGEPKIIPEKNSLRQTYNSCARLCLNQETVCLA 60

61 STAMKTEVCVAKTKLANGTSSMIVPKQKLSASYEKEKELCVKYPQWSESQDQVEFVHL 120

61 STAMKTEVCVAKTKLANGTSSMIVPKQKLSASYEKEKELCVKYPQWSESQDQVEFVHL 120

121 ISOMCHYQHGHSINLYKPMQLQDFTALPARGLDHIAENILSYLDAKSLCAELNCKEY 180

121 ISOMCHYQHGHSINLYKPMQLQDFTALPARGLDHIAENILSYLDAKSLCAELNCKEY 180

181 RVTSDGMLWKKLIERMVRTDSLWRGLAERRGWQYLFKNKPPDGNAPPNSFYRALYPKII 240

181 RVTSDGMLWKKLIERMVRTDSLWRGLAERRGWQYLFKNKPPDGNAPPNSFYRALYPKII 240

241 QDIETIESNWRGCRHSQRICRSETSGKVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300

241 QDIETIESNWRGCRHSQRICRSETSGKVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300

301 RILTGHTGSLVLCQYDERVITGSSDSTVRVWDVNTGEMNTLIHCEAVLHFRNNGMW 360

301 RILTGHTGSLVLCQYDERVITGSSDSTVRVWDVNTGEMNTLIHCEAVLHFRNNGMW 360

361 VTCSDKRSIAVMDNASPTDITLRRVLVGHRAAVNVVDFDKYIVSASGDRTIKWNNTSTC 420

361 VTCSDKRSIAVMDNASPTDITLRRVLVGHRAAVNVVDFDKYIVSASGDRTIKWNNTSTC 420

421 EFPVTLNGHKGIAQLQYDRDLVYSGSDNTIRLWDETCGACLRVLEGGHEELVRCIRFDN 480

421 EFPVTLNGHKGIAQLQYDRDLVYSGSDNTIRLWDETCGACLRVLEGGHEELVRCIRFDN 480

481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTCLRLTVLHSGRVRFLQDFEQIVSSSHDDT 540

481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTCLRLTVLHSGRVRFLQDFEQIVSSSHDDT 540

541 ILIWFNDPAAQAEPPSPRTYIISR 569

541 ILIWFNDPAAQAEPPSPRTYIISR 569

7

322446 standard; Protein; 569 AA.

322446;

-OCT-2002 (first entry)

XX DE Human F-box protein FBp1 SEQ ID No 2.

XX KW Cytostatic; immunomodulator; cardiant; antiinflammatory; antimicrobial;
XX KW proliferative; differentiative disorder; Skp2; F-box protein; cancer;
XX KW ubiquitin ligase; breast cancer; prostate cancer; ovarian cancer;
XX KW small cell lung carcinoma; immune disorder; parathyroid adenoma; FBP;
XX KW inflammatory disorder; lymphoma; major opportunistic infection;
XX KW certain cardiovascular disease; human.

XX OS Homo sapiens.

XX PN WO200255665-A2.

XX PD 18-JUL-2002.

XX PF 07-JAN-2002; 2002WO-US00311.

XX PR 05-JAN-2001; 2001US-260179P.

XX PA (UJNY) UNIV NEW YORK STATE.

XX Pagano M;

XX DR MPI; 2002-599665/64.

XX DR N-PSDB; AAL41041.

XX PT Screening compounds for treating proliferative disorders, e.g. breast
XX PT cancer or prostate cancer, infections or immune disorders, comprises
XX PT detecting a change in the activity of Skp2 with either p27 or Cks1 -

XX PS Disclosure; Fig 3; 246pp; English.

XX CC The invention relates to screening compounds useful for the treatment of
XX CC proliferative or differentiative disorders comprising detecting a change
XX CC in the activity of Skp2 (F-box protein). The method is useful for
XX CC screening compounds for the treatment of proliferative or differentiative
XX CC disorders, particularly cancer. These compounds include small molecules,
XX CC or compounds or derivatives or analogues of the new ubiquitin ligases.
XX CC The compounds are useful for treating diseases such as cancer (e.g.
XX CC breast cancer, prostate cancer or ovarian cancer, lymphoma, small cell
XX CC lung carcinoma or parathyroid adenomas), major opportunistic infections,
XX CC immune disorders, certain cardiovascular diseases or inflammatory
XX CC disorders. This sequence represents an F-box protein (FBP) relating to
XX CC the invention.

XX SQ Sequence 569 AA;

Query Match 100.0%; Score 3034; DB 23; Length 569;
Best Local Similarity 100.0%; Pred. No. 2.5e-286;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAEAVLOEKALFKMNSEREDCNNGEPKIIPEKNSLRQTYNSCARLCLNQETVCLA 60

Db 1 MDPAEAVLOEKALFKMNSEREDCNNGEPKIIPEKNSLRQTYNSCARLCLNQETVCLA 60

QY 61 STAMKTEVCVAKTKLANGTSSMIVPKQKLSASYEKEKELCVKYPQWSESQDQVEFVHL 120

Db 61 STAMKTEVCVAKTKLANGTSSMIVPKQKLSASYEKEKELCVKYPQWSESQDQVEFVHL 120

QY 121 ISOMCHYQHGHSINLYKPMQLQDFTALPARGLDHIAENILSYLDAKSLCAELNCKEY 180

Db 121 ISOMCHYQHGHSINLYKPMQLQDFTALPARGLDHIAENILSYLDAKSLCAELNCKEY 180

QY 181 RVTSDGMLWKKLIERMVRTDSLWRGLAERRGWQYLFKNKPPDGNAPPNSFYRALYPKII 240

Db 181 RVTSDGMLWKKLIERMVRTDSLWRGLAERRGWQYLFKNKPPDGNAPPNSFYRALYPKII 240

QY 241 QDIETIESNWRGCRHSQRICRSETSGKVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300

Db 241 QDIETIESNWRGCRHSQRICRSETSGKVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300

QY 301 RILTGHTGSLVLCQYDERVITGSSDSTVRVWDVNTGEMNTLIHCEAVLHFRNNGMW 360

```

301 RLITGHTGVLCLQYDERVITITGSSDSTVRVMDVNTGEMLNTLIHCEAVLHFRNNGMM 360
361 VTCSKDRSIAVWDMASPTDITLRLVLVGHRAAVNVVDFDDKYIVSASGDRITIKWNTSTC 420
361 VTCSKDRSIAVWDMASPTDITLRLVLVGHRAAVNVVDFDDKYIVSASGDRITIKWNTSTC 420
421 EFVRTLNGHKGRIACIQYRDLRVVSGSSDNTIIRLWDIECCACLRVLEGEHELVRGIRFDN 480
421 EFVRTLNGHKGRIACIQYRDLRVVSGSSDNTIIRLWDIECCACLRVLEGEHELVRGIRFDN 480
481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTLVEHSGRVFRLOQDFEQIVSSSHDDT 540
481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTLVEHSGRVFRLOQDFEQIVSSSHDDT 540
541 ILIWFNDPAAQAEPSPSRRTTYISR 569
541 ILIWFNDPAAQAEPSPSRRTTYISR 569

G69473 standard; Protein; 569 AA.
G69473;
-OCT-2002 (first entry)
man bait protein beta-TrCP1.
man; yeast two-hybrid assay; adipocyte; bait protein; NIDDM;
n-insulin diabetes mellitus; obesity; selected interacting domain;
D; protein-protein interaction map; PIM; anorectic; metabolic disorder.
mo sapiens.
200253726-A2.
-JUL-2002.
-DEC-2001; 2001WO-EP15423.
-JAN-2001; 2001US-259377P.
YBR-} HYBRIGENICS.
YRS } CENT NAT RECH SCI.
grain P, Marullo S, Jockers R;
I; 2002-583612/62.
>SDB; ABS51009.
vel complex of protein-protein interactions in adipocyte cells for
antifungal compounds that modulate the protein-protein interactions
a useful for treating obesity and metabolic disorders -
aim 1; Page -: 125pp; English.
a invention relates to a complex of protein-protein interactions
arming a protein-protein interaction map, PIM) in adipocyte cells as
fined in the specification, or polynucleotides in adipocytes encoding
r the polypeptides. Also included are a recombinant cell expressing the
eracting polypeptides and a method of selecting a modulating compound
adipocyte cells, by cultivating a recombinant host cell on a selective
tium containing a modulating compound and a reporter gene the
pression of which is toxic for the recombinant host cell which is
nsformed with two vectors, where the first vector comprises a
ynucleotide encoding a first hybrid polypeptide and DNA binding
ain and the second vector comprising a polynucleotide encoding a
ond hybrid polypeptide and an activating domain that activates the
tic reporter gene, when the first and second hybrid polypeptides
eract and selecting the modulating compound which inhibits the
wth of the recombinant host cell (i.e. using the yeast two-hybrid

```

```

CC system). The complexes are useful for identifying compounds that modulate
CC the protein-protein interactions and useful for treating obesity and
CC metabolic disorders e.g. non-insulin dependent diabetes mellitus,
CC NIDDM. The compound isolated by the method is useful for treating and
CC preventing obesity or metabolic diseases. The interactions between
CC the proteins of the complex further define a set of selected interacting
CC domains, SIP. The present sequence represents a member of the protein
CC complex of the invention, used as the bait protein in the yeast two-
CC hybrid assay.
CC Note: The present sequence was not displayed in the specification but
CC was obtained from its Genbank entry by the indexer.
XX
SQ Sequence 569 AA;
Query Match 100.0%; Score 3034; DB 23; Length 569;
Best Local Similarity 100.0%; Pred. No. 2.5e-286;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDPAAVLOEKALKFNMSSEREDCNNGEPKRIIEKNSLRQTYNSCARCLNQETVCLA 60
DB 1 MDPAAVLOEKALKFNMSSEREDCNNGEPKRIIEKNSLRQTYNSCARCLNQETVCLA 60
QY 61 STAMKTENCVAKTKLANGTSSMIVPKQKLSASYEKEKELCVKYPEQWSESQVDFVEHL 120
DB 61 STAMKTENCVAKTKLANGTSSMIVPKQKLSASYEKEKELCVKYPEQWSESQVDFVEHL 120
QY 121 ISQWCHYQHGHNISYLPMLQDFITALPARGLDHIAENILSYLDAKSLCAAEVCKEWY 180
DB 121 ISQWCHYQHGHNISYLPMLQDFITALPARGLDHIAENILSYLDAKSLCAAEVCKEWY 180
QY 181 RVTSQGLMKKLLIERVWRTDSLWRGLAERRGGQVLFKNKPPDGNAPPNSFYRALYPKII 240
DB 181 RVTSQGLMKKLLIERVWRTDSLWRGLAERRGGQVLFKNKPPDGNAPPNSFYRALYPKII 240
QY 241 QDIETTESNWRCCGRHSIQIHCRSETSKGVYCLQYDDQKIVSGLRDNITIKWDKNTLECK 300
DB 241 QDIETTESNWRCCGRHSIQIHCRSETSKGVYCLQYDDQKIVSGLRDNITIKWDKNTLECK 300
QY 301 RLITGHTGVLCLQYDERVITITGSSDSTVRVMDVNTGEMLNTLIHCEAVLHFRNNGMM 360
DB 301 RLITGHTGVLCLQYDERVITITGSSDSTVRVMDVNTGEMLNTLIHCEAVLHFRNNGMM 360
QY 361 VTCSKDRSIAVWDMASPTDITLRLVLVGHRAAVNVVDFDDKYIVSASGDRITIKWNTSTC 420
DB 361 VTCSKDRSIAVWDMASPTDITLRLVLVGHRAAVNVVDFDDKYIVSASGDRITIKWNTSTC 420
QY 421 EFVRTLNGHKGRIACIQYRDLRVVSGSSDNTIIRLWDIECCACLRVLEGEHELVRGIRFDN 480
DB 421 EFVRTLNGHKGRIACIQYRDLRVVSGSSDNTIIRLWDIECCACLRVLEGEHELVRGIRFDN 480
QY 481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTLVEHSGRVFRLOQDFEQIVSSSHDDT 540
DB 481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTLVEHSGRVFRLOQDFEQIVSSSHDDT 540
QY 541 ILIWFNDPAAQAEPSPSRRTTYISR 569
DB 541 ILIWFNDPAAQAEPSPSRRTTYISR 569
RESULT 9
ID AAB48298 standard; protein; 569 AA.
XX AAB48298;
AC AAB48298;
XX
DT 02-APR-2001 (first entry)
XX Human ZF11 protein.
XX
KW S-phase kinase associated protein; SKP1; SKP2; SKP2-like protein; ZF;
KW CUL-1; cullin; CDC53; p27; cyclin E; Max; Mad; c-Myc; MDM2; p53; Bax;
KW Bad; Bcl-2; tumour; cytostatic.
XX

```

omo sapiens.
D200075184-A1.
4-DEC-2000.
5-JUN-2000; 2000WO-US15449.
1-JUN-1999; 99US-0137494.
JYYA) UNIV YALE.
iang H, Tsvetkov IM, Kondo T;
PI; 2001-061703/07.
-PSDB; AAC84610.
odulating polypeptide levels in a cell, diagnosing and treating tumor,
volves altering levels of proteins such as S-phase kinase associated
roteins 1, 2 and cullin/CDC53 proteins -
laim 3; Page 130-132; 162pp; English.
e invention relates to methods of altering the polypeptide levels in a
ll, using proteins selected from S-phase kinase associated proteins 1
d 2 (SKP1, SKP2), SKP2-like proteins (ZF) and CUL-1 (a member of the
illin/ CDC53 family of proteins). The method is useful for altering the
vel of p27, cyclin E, Max, Mad, c-Myc, MDM2, p53, Bax, Bad or Bcl-2
ypeptide in a cell. SKP2 and SKP2-like protein levels are useful for
ecting tumours, and in monitoring tumor treatment in a mammal. Agents
at modulate interactions between SKP and target proteins are useful for
eating tumours.
quence 569 AA;
Match 99.8%; Score 3027; DB 22; Length 569;
Local Similarity 99.8%; Pred. No. 1.2e-285;
es 568; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 MDPAEVLQEKALFPMNSERDCNNGEPKPKIPEKNSLRQTVNSCARLCLNQETVCLA 60
1 MDPAEVLQEKALFPMNSERDCNNGEPKPKIPEKNSLRQTVNSCARLCLNQETVCLA 60
61 STAMKTECNVAKTKLANGTSMIVPKOKLSASVEKEKELCVKVFQWSSDOVEFVEHL 120
61 STAMKTECNVAKTKLANGTSMIVPKOKLSASVEKEKELCVKVFQWSSDOVEFVEHL 120
121 ISQWCHYQGHGINSYLPKMLQDFITLAPARGLDHIATENILSYLDAKSLCAAEVLCKEWY 180
121 ISQWCHYQGHGINSYLPKMLQDFITLAPARGLDHIATENILSYLDAKSLCAAEVLCKEWY 180
181 RVTSDGMLWKKLIERWRTDSLRGLAERRGWGYLFKNKPPDGNAPPNSFYALYPKII 240
181 RVTSDGMLWKKLIERWRTDSLRGLAERRGWGYLFKNKPPDGNAPPNSFYALYPKII 240
241 QDIETIESNRCGRHSIORHCRSETSGKVYCLQYDQKIVSGLRDNTIKWDKNTLECK 300
241 QDIETIESNRCGRHSIORHCRSETSGKVYCLQYDQKIVSGLRDNTIKWDKNTLECK 300
301 RLITGHTGSLVCLQYDERVITGSSDSTVRVWDNTGEMTLNLIHCEAVLHLPFNGWM 360
301 RLITGHTGSLVCLQYDGRVITGSSDSTVRVWDNTGEMTLNLIHCEAVLHLPFNGWM 360
361 VTCSKDRSIADVMASTPDIITLRLVYLGHRAAVNVVDFDKYIVSASGDRITIKWNTSTC 420
361 VTCSKDRSIADVMASTPDIITLRLVYLGHRAAVNVVDFDKYIVSASGDRITIKWNTSTC 420
421 EFVRTLNGHKRGIAQLQVDRVLVSGSSDNTIRLWDIECGACLRVLGHEELVRCIRFDN 480
421 EFVRTLNGHKRGIAQLQVDRVLVSGSSDNTIRLWDIECGACLRVLGHEELVRCIRFDN 480
481 KRIVSGAYDGKIKVWDVAALDPRAPAGTLCRLTVEHSGRVRFLQDFQIVSSSHDDT 540

Db 481 KRIVSGAYDGKIKVWDVAALDPRAPAGTLCRLTVEHSGRVRFLQDFQIVSSSHDDT 540
Qy 541 ILIWDFLNDPAQAQAEPPRSPRSTYTYISR 569
Db 541 ILIWDFLNDPAQAQAEPPRSPRSTYTYISR 569
RESULT 10
AAM00960
ID AAM00960 standard; Protein; 608 AA.
XX
AC AAM00960;
XX
DT 01-OCT-2001 (first entry)
XX
DE Human bone marrow protein, SEQ ID NO: 436.
XX
KW Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
KW immunosuppressive; gene therapy; cytokine cell proliferation;
KW cell differentiation modulator; immune disorder; infection; cancer;
KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.
XX
OS Homo sapiens.
XX
FN WO200153453-A2.
XX
PD 26-JUL-2001.
XX
PF 23-DEC-2000; 2000WO-US34960.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 30-NOV-2000; 2000US-0250593.
XX
PA (HYSE-) HYSE INC.
XX
PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Drmanac RT;
XX
DR WPI; 2001-488707/53.
DR N-PSDB; AAH90079.
XX
PT Novel bone-marrow-expressed polynucleotides and polypeptides, useful
for treating e.g. cancer and immune deficiency disorders -
XX
PS Claim 10; Page 523-524; 648pp; English.
XX
CC The present sequence is one of 251 novel human polypeptides encoded
by a bone marrow-expressed polynucleotide. The polynucleotide and the
polypeptide encoded by it are useful in the treatment of various
immune deficiencies and disorders. The deficiencies and disorders may
be genetic, may be caused by a viral (e.g. HIV) bacterial or fungal
infection, or may result from an autoimmune disorder, a coagulation
disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
suppression of an inflammatory response or treatment of a nervous
system disorder such as Alzheimer's disease. Detection of the presence
or increased expression of the polynucleotide or the protein it
encodes is useful for the diagnosis and/or prognosis of one
or more types of cancer. The polynucleotide and polypeptide can be
used as nutritional sources or supplements and in the screening of
chemical compounds as potential drugs.
XX
XX Sequence 608 AA;
Query Match 99.8%; Score 3027; DB 22; Length 608;
Best Local Similarity 99.6%; Pred. No. 1.4e-285;


```
tes 567; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
1 MDPAEVLQEKALFPMNSREDCCNGEPPRKIIPEKNSLRQTYNSCARLCLNQETVCLA 60
40 MDPVEAVLQEKALFPMNSREDCCNGEPPRKIIPEKNSLRQTYNSCARLCLNQETVCLA 99
61 STAMKTENCVAKTCLANGTSSMIVPKOKLSASVEKEKELCVKYPEQWSESDDQVEVEHL 120
100 STAMKTENCVAKTCLANGTSSMIVPKOKLSASVEKEKELCVKYPEQWSESDDQVEVEHL 159
121 ISQCHYQGHGHSYLYKPMQLQDFITLARGLDHIAENILSYLDAKSLCAAEVLCKEY 180
160 ISQCHYQGHGHSYLYKPMQLQDFITLARGLDHIAENILSYLDAKSLCAAEVLCKEY 219
181 RVSDGMLWKKLIERMVRTDSLWGLAERRGQGYLFPKNPPDGNAPPNSFYRALKPKII 240
220 RVSDGMLWKKLIERMVRTDSLWGLAERRGQGYLFPKNPPDGNAPPNSFYRALKPKII 279
241 QDIETIESNRCGRHSQRHICRSETSGVYCYQYDDQKIVSGLRONTIKINDNTLECK 300
280 QDIETIESNRCGRHSQRHICRSETSGVYCYQYDDQKIVSGLRONTIKINDNTLECK 339
301 RILGHTGSVLCLOYDERVIITGSDSTVRVWVNTGEMLNTLIHCEAVLHURFNNGMM 360
340 RILGHTGSVLCLOYDERVIITGSDSTVRVWVNTGEMLNTLIHCEAVLHURFNNGMM 399
361 VTCSDRSIAVMDVASTDITLRLVGHRAAVNVVDFDDKIVSASGDRTIKVMNTSTC 420
400 VTCSDRSIAVMDVASTDITLRLVGHRAAVNVVDFDDKIVSASGDRTIKVMNTSTC 459
421 EFVTLNGHKGRIACLOVRDLVVGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN 480
460 EFVTLNGHKGRIACLOVRDLVVGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN 519
481 KRIVSGAYDGIKYWDLVAALDPRAPAGTLCRLTLVEHSGRVFRLQDFEQIVSSSHDDT 540
520 KRIVSGAYDGIKYWDLVAALDPRAPAGTLCRLTLVEHSGRVFRLQDFEQIVSSSHDDT 579
541 ILIWDFLNDPAAQAEPPRSPRTYIISR 569
580 ILIWDFLNDPAAQAEPPRSPRTYIISR 608
11
2
M78582 standard; Protein; 605 AA.
-NOV-2001 (first entry)
man protein SEQ ID NO 1244.
man; cytokine; cell proliferation; cell differentiation; gene therapy;
ccine; peptide therapy; stem cell growth factor; haematopoiesis;
issue growth factor; immunomodulatory; cancer; leukaemia;
rvous system disorder; arthritis; inflammation.
mo sapiens.
200157190-A2.
-AUG-2001.
-FEB-2001; 2001WO-0504098.
-FEB-2000; 2000US-0496914.
-APR-2000; 2000US-0560875.
-JUN-2000; 2000US-0598075.
-JUL-2000; 2000US-0620325.
-SEP-2000; 2000US-0654936.
-SEP-2000; 2000US-0663561.
-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX (HYSE-) HYSEQ INC.
PA Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
XX Zhao Qa, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX WPI; 2001-476283/51.
DR N-PSDB; AAK51715.
XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX Claim 20; Page 3503-3504; 6221pp; English.
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAK78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAK80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX SQ Sequence 605 AA;
Query Match 99.1%; Score 3006; DB 22; Length 605;
Best Local Similarity 94.0%; Pred. No. 1.5e-283;
Matches 569; Conservative 0; Mismatches 0; Indels 36; Gaps 1;
QY 1 MDPAEVLQEKALK-----FMNSREDCC 24
DB 1 MDPAEVLQEKALKFMCMSPRSLWLGSSLDASMPSLRCLYNGTGALTAFMNSREDCC 60
QY 25 NNGEPPRKIIPEKNSLRQTYNSCARLCLNQETVCLASTAMKTENCVAKTCLANGTSSMIV 84
DB 61 NNGEPPRKIIPEKNSLRQTYNSCARLCLNQETVCLASTAMKTENCVAKTCLANGTSSMIV 120
QY 85 PKOKLSASVEKEKELCVKYPEQWSESDDQVEVEHLISQCHYQGHGHSYLYKPMQLQDF 144
DB 121 PKOKLSASVEKEKELCVKYPEQWSESDDQVEVEHLISQCHYQGHGHSYLYKPMQLQDF 180
QY 145 ITALPARGLDHIAENILSYLDAKSLCAAEVLVCKEYRVTSDGMLWKKLIERMVRTDSLWR 204
DB 181 ITALPARGLDHIAENILSYLDAKSLCAAEVLVCKEYRVTSDGMLWKKLIERMVRTDSLWR 240
QY 205 GLAERRGQGYLFPKNPPDGNAPPNSFYRALKPKIIQDIETIESNRCGRHSQRHICRS 264
DB 241 GLAERRGQGYLFPKNPPDGNAPPNSFYRALKPKIIQDIETIESNRCGRHSQRHICRS 300
QY 265 ETSKGVYCYQYDDQKIVSGLRDNTIKINDNTLECKRILTGHTGSLVLCLOYDERVIITGS 324
DB 301 ETSKGVYCYQYDDQKIVSGLRDNTIKINDNTLECKRILTGHTGSLVLCLOYDERVIITGS 360
QY 325 SDSTVRVWVNTGEMLNTLIHCEAVLHURFNNGMMVTCSDRSIAVMDVASTDITLRR 384
DB 361 SDSTVRVWVNTGEMLNTLIHCEAVLHURFNNGMMVTCSDRSIAVMDVASTDITLRR 420
QY 385 VLVGHRAAVNVVDPDDKIVSASGDRTIKVMNTSTCEFVTLNGHKGRIACLOVRDLVV 444
DB 421 VLVGHRAAVNVVDPDDKIVSASGDRTIKVMNTSTCEFVTLNGHKGRIACLOVRDLVV 480
QY 445 SGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDNKRIVSGAYDGIKYWDLVAALDPR 504
DB 481 SGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDNKRIVSGAYDGIKYWDLVAALDPR 540
```

505 APAGTLCRLTVEHSGRVFRLQDFEQIVSSSHDDTILIMDFLNDPAAQAEPPSPRTY 564
|||||
541 APAGTLCRLTVEHSGRVFRLQDFEQIVSSSHDDTILIMDFLNDPAAQAEPPSPRTY 600
|||||
565 TYISR 569
|||||
601 TYISR 605
|||||
12
12
AB12812 standard; protein; 569 AA.
AB12812;
7-NOV-2000 (first entry)
ouse ubiquitin ligase FWD1 protein SEQ ID NO:2.
ubiquitin ligase SCF complex; F-box protein; ubiquitination; IkappaB;
ata-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FWD1;
ne therapy; colon cancer; beta-transducin repeat containing protein;
ata-TrCP.
is musculus.
22000166542-A.
)-JUN-2000.
-DEC-1998; 98JP-0343437.
-DEC-1998; 98JP-0343437.
GAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
I; 2000-485550/43.
PSDB; AAA73131.
box protein of ubiquitin ligase SCF complex which promotes the
ubiquitination of IkappaB or Beta-catenin -
aim 2; Page 9-10; 19pp; Japanese.
e present invention describes an F-box motif protein of ubiquitin
gase SCF complex which promotes the ubiquitination of IkappaB or
ta-catenin and is constituted by Skp1 protein, Cull1 protein and a
mplex (SCF complex) of F-box protein containing F-box motif and a
peat motif and has the amino acid sequence of 45 residues (AAB12811)
one of two 569 residue sequences (AAB12812, which is mouse ubiquitin
gase FWD1 protein) and (AAB12813, which is human beta-transducin
peat containing protein (beta-TrCP)). The F-box protein can be used for
e gene therapy of colon cancer by being recombined to a virus vector.
quence 569 AA;
Match
Local Similarity 98.8%; Score 2997; DB 21; Length 569;
es 561; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
1 MDPAAVLQEKALKFNNSSRDCCNGGPPPKIIPKNSLRQTYNSCARLCLNQETVCLA 60
1 MDPAAVLQEKALKFNNSSRDCCNGGPPPKIIPKNSLRQTYNSCARLCLNQETVCLT 60
61 STAMKTENCVAKTLKANGTSSMIVPKQKLSASYEKEXELCVKVFQWSESDDQVFVEHL 120
61 STAMKTENCVAKALANGTSSMIVPKQKLSASYEKEXELCVKVFQWSESDDQVFVEHL 120
121 ISQCHYQHGHNINSYKPLMQDFITALPARGLDHIAENILSYLDKSLCAAEILVCKEWY 180
121 ISQCHYQHGHNINSYKPLMQDFITALPARGLDHIAENILSYLDKSLCAAEILVCKEWY 180
181 RVTSDGMLWKKLIERMVRTDSLWRGLAERRGWSQYLFKNKPPDENAPPNSFYRALYPKII 240

DB 181 RVTSDGMLWKKLIERMVRTDSLWRGLAERRGWSQYLFKNKPPDENAPPNSFYRALYPKII 240
QY 241 ODIEETISNWRGHSIORIHCSETSKGVYCLQYDDOKIVSGRLDNTIKIWDKNTLECK 300
DB 241 ODIEETISNWRGHSIORIHCSETSKGVYCLQYDDOKIVSGRLDNTIKIWDKNTLECK 300
QY 301 RILTGHTGSLVCLQYDERVIIITGSSDSTVRVNDVNTGEMINTLIHCEAVLHLRFNNGMM 360
DB 301 RILTGHTGSLVCLQYDERVIIITGSSDSTVRVNDVNTGEMINTLIHCEAVLHLRFNNGMM 360
QY 361 VTCSDRSIAWDMASPTDITLRLVLVGHRAAANNVDFDDKYIVSASGDRTIKWNNTSTC 420
DB 361 VTCSDRSIAWDMASPTDITLRLVLVGHRAAANNVDFDDKYIVSASGDRTIKWNNTSTC 420
QY 421 EFVETLNGHKGKGIACQYRDLRVVSGSSDNTIRLWDIECGACLRVLEGEHELVRGIRFDN 480
DB 421 EFVETLNGHKGKGIACQYRDLRVVSGSSDNTIRLWDIECGACLRVLEGEHELVRGIRFDN 480
QY 481 KRIVSGAYDGKIKVMDLVAAALDPRAPAGTLCRLTVEHSGRVFRLQDFEQIVSSSHDDT 540
DB 481 KRIVSGAYDGKIKVMDLVAAALDPRAPAGTLCRLTVEHSGRVFRLQDFEQIVSSSHDDT 540
QY 541 ILIWDFLNDPAAQAEPPSPRTYTYISR 569
DB 541 ILIWDFLNDPAAQAEPPSPRTYTYISR 569
RESULT 13
AAV83254
ID AAV83254 standard; Protein; 569 AA.
XX
AC AAV83254;
XX
DT 16-AUG-2000 (first entry)
XX
XX F-box protein FWD1p.
XX Ubiquitin ligase; SCF; F-box protein; targeted degradation;
KW destabilisation; proteolysis; drug discovery; gene therapy; cancer;
KW oncoprotein; Huntington's disease; gene knockout; delivery systems;
KW mouse; ss.
XX
OS Mus musculus.
XX
XX WO2000022110-A2.
XX
XX 20-APR-2000.
XX 08-OCT-1999; 99WO-US23705.
XX 09-OCT-1998; 98US-0103787.
XX (HARD) HARVARD COLLEGE.
XX
XX Zhou P, Howley P;
PI
XX WPI; 2000-317970/27.
DR N-PSDB; AMZ93714.
XX
XX Targeting degradation of polypeptide useful for treating cancer and
PT other proliferative disorders, involves conjugating polypeptide with
PT ubiquitin protein ligase or inhibiting ubiquitination using organic
PT compound
XX
PS Claim 9; Page 184-185; 185pp; English.
XX
CC The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin
CC ligases) which can be used for the targeted degradation of a target
CC polypeptide in vivo. Targeted degradation is achieved by expressing
CC the ubiquitin ligase in a cell linked to the interaction domain of
CC the target polypeptide and thereby recruiting the target polypeptide
CC to the ubiquitin ligase. Such methods are useful for decreasing or

increasing the level of a target polypeptide and for creating and expressing a destabilized polypeptide which is subjected to SCF mediated proteolysis. Degrading any desired protein in a cell is useful for preventing or treating diseases caused by the presence of abnormal amount of the specific polypeptides, for drug discovery and or gene therapy. Diseases treated include cancer, by degradation of oncoproteins, Huntington's disease, other proliferative disorders and microbial infections. The method provides a quick and easy alternative to gene knockout technology. The target polypeptide can be degraded at all stages, or a specific stage, of development in the mature animal.

sequence 569 AA;
Y Match 98.8%; Score 2997; DB 21; Length 569;
Local Similarity 98.6%; Pred. No. 1e-282;
hes 561; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

1 MDPAAVLQEKALFPMNSEREDCNNGEPKIIPEKNSLRQTYNSCARLCLNQETVCLA 60
1 MDPAAVLQEKALFPMNSEREDCNNGEPKIIPEKNSLRQTYNSCARLCLNQETVCLT 60
61 STAMKTCNCVAKTKLANGTSSMIVPKOKLSASVEKEKELCVKYFEOWSSDQVEFV 120
61 STAMKTCNCVAKTKLANGTSSMIVPKOKLSASVEKEKELCVKYFEOWSSDQVEFV 120
121 ISQMCHYQGHINSYKPMQLQDFITLALPARGLDHIAENILSYLDAKSLCAELVCKEW 180
121 ISQMCHYQGHINSYKPMQLQDFITLALPARGLDHIAENILSYLDAKSLCAELVCKEW 180
181 RVTSQGLMVKLLIERMVRTDLSLWGLAERRGWGYLFKNKPPDGNAPPNSFYALPKII 240
181 RVTSQGLMVKLLIERMVRTDLSLWGLAERRGWGYLFKNKPPDGNAPPNSFYALPKII 240
241 QDIETIESNWRGCRHSLOIHCRTSETSGVYCLQYDDQKIVSGLRDNITIKWDKNTLECK 300
241 QDIETIESNWRGCRHSLOIHCRTSETSGVYCLQYDDQKIVSGLRDNITIKWDKNTLECK 300
301 RILHTGTSVLCLOYDERVITIGSSDSTVRVWVNTGEMNTLIHHCCEAVLHFRFNGWM 360
301 RILHTGTSVLCLOYDERVITIGSSDSTVRVWVNTGEMNTLIHHCCEAVLHFRFNGWM 360
361 VTCKORSIAVNDWASPTDITLRLVGHRAAVNVDFDKYIVSASGDTIKVWNTSTC 420
361 VTCKORSIAVNDWASPTDITLRLVGHRAAVNVDFDKYIVSASGDTIKVWNTSTC 420
421 EFVRTLNGHKGRIACLOYRDLVWVGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN 480
421 EFVRTLNGHKGRIACLOYRDLVWVGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN 480
481 KRIVSGAYDGKIKWDLVAALDPRAPAGTLCRLTLVEHSGRVFRLQDFEQIVSSSHDDT 540
481 KRIVSGAYDGKIKWDLVAALDPRAPAGTLCRLTLVEHSGRVFRLQDFEQIVSSSHDDT 540
541 ILIWDFLNDPAAQEPSPSRITYISR 569
541 ILIWDFLNDPAAHAEPPSPSRITYISR 569

14
4

M78584 standard; Protein; 632 AA.

M78584;

-NOV-2001 (first entry)

nan protein SEQ ID NO 1246.

nan; cytokine; cell proliferation; cell differentiation; gene therapy;
cine; peptide therapy; stem cell growth factor; haematopoiesis;
sue growth factor; immunomodulatory; cancer; leukaemia;
rvous system disorder; arthritis; inflammation.

XX OS Homo sapiens.
XX PN WO200157190-A2.
XX PD 09-AUG-2001.
XX PF 05-FEB-2001; 2001WO-US040398.
XX PR 03-FEB-2000; 2000US-0496914.
XX PR 27-APR-2000; 2000US-0560875.
XX PR 20-JUN-2000; 2000US-0598075.
XX PR 19-JUL-2000; 2000US-0620325.
XX PR 01-SEP-2000; 2000US-0654936.
XX PR 15-SEP-2000; 2000US-0663541.
XX PR 20-OCT-2000; 2000US-0693325.
XX PR 30-NOV-2000; 2000US-0728422.
XX (HYSE-) HYSEQ INC.
XX PA Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
XX PI Zhao QA, Wang D, Zhang J, Ren F, Chen R, Wang ZW;
XX PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX DR WPI; 2001-476283/51.
XX DR N-PSDB; AAK51717.
XX Nucleic acids encoding polypeptides with cytokine-like activities,
XX useful in diagnosis and gene therapy -
XX Claim 20; Page 3505-3507; 6221pp; English.
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation.
XX Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
XX (AAM80020) are omitted as the relevant pages from the sequence listing
XX were missing at the time of publication.

SQ Sequence 632 AA;

Query Match 98.6%; Score 2992.5; DB 22; Length 632;
Best Local Similarity 90.0%; Pred. No. 3.3e-282;
Matches 569; Conservative 0; Mismatches 0; Indels 63; Gaps 1;

QY 1 MDPAAVLQEKALK-----FMNSEREDCNNGEPKIIPEKNSLRQTYNSCARLCLNQETV 14
DB 1 MDPAAVLQEKALKFMFMEPRSCGWNMTWARSRLTATSTSRVQCSMPERSLWLGSSSLADS 60
QY 15 -----FMNSEREDCNNGEPKIIPEKNSLRQTYNSCARLCLNQETV 57
DB 61 MPELRCLYNPGTGTALTAFMNSEREDCNNGEPKIIPEKNSLRQTYNSCARLCLNQETV 120
QY 58 CLASTAMKTCNCVAKTKLANGTSSMIVPKOKLSASVEKEKELCVKYFEOWSSDQVEFV 117
DB 121 CLASTAMKTCNCVAKTKLANGTSSMIVPKOKLSASVEKEKELCVKYFEOWSSDQVEFV 180
QY 118 EHLISQMCHYQGHINSYKPMQLQDFITLALPARGLDHIAENILSYLDAKSLCAELVCK 177
DB 181 EHLISQMCHYQGHINSYKPMQLQDFITLALPARGLDHIAENILSYLDAKSLCAELVCK 240
QY 178 EYRVTSQGLMVKLLIERMVRTDLSLWGLAERRGWGYLFKNKPPDGNAPPNSFYALYP 237
DB 241 EYRVTSQGLMVKLLIERMVRTDLSLWGLAERRGWGYLFKNKPPDGNAPPNSFYALYP 300

238 KIIQDIETIESNRCGRHSLSQIHCRTSSTSGVYCLQYDDQKIVSGLRDNTIKIWDKNTL 297
 |||||
 301 KIIQDIETIESNRCGRHSLSQIHCRTSSTSGVYCLQYDDQKIVSGLRDNTIKIWDKNTL 360
 |||||
 298 ECKRIILTGTGTVLCLQYDERVLIITGSSDSTVRVWDVNTGEMLNTLIHHCCEAVLHLPFNN 357
 |||||
 361 ECKRIILTGTGTVLCLQYDERVLIITGSSDSTVRVWDVNTGEMLNTLIHHCCEAVLHLPFNN 420
 |||||
 358 GMMVTCSKDRSIAVMDASPTDITLRLVGVHRAAVNVVDFDDKIYIVSASGDRTIKWNT 417
 |||||
 421 GMMVTCSKDRSIAVMDASPTDITLRLVGVHRAAVNVVDFDDKIYIVSASGDRTIKWNT 480
 |||||
 418 STCEFVRLTNGHKGRIACLOVEDRLVSGSSDNTIRLWDIECGACLRVLEGHEELVRCIR 477
 |||||
 481 STCEFVRLTNGHKGRIACLOVEDRLVSGSSDNTIRLWDIECGACLRVLEGHEELVRCIR 540
 |||||
 478 FDNKRIVSGAYDGKIKVMDLVAALDPAPAGTLCRLTLVEHSGRVFRLQDFEFOIVSSSH 537
 |||||
 541 FDNKRIVSGAYDGKIKVMDLVAALDPAPAGTLCRLTLVEHSGRVFRLQDFEFOIVSSSH 600
 |||||
 538 DDTILWDFLNDPAAQAEPPSPSRTYIYSR 569
 |||||
 601 DDTILWDFLNDPAAQAEPPSPSRTYIYSR 632
 |||||

15
 56

AM79566 standard; Protein; 654 AA.

AM79566;

5-NOV-2001 (first entry)

uman protein SEQ ID NO 3212.

uman; cytokine; cell proliferation; cell differentiation; gene therapy;
 acine; peptide therapy; stem cell growth factor; haematopoiesis;
 issue growth factor; immunomodulatory; cancer; leukaemia;
 vious system disorder; arthritis; inflammation.

mo sapiens.

Y200157190-A2.

-AUG-2001.

-FEB-2001; 2001WO-US04098.

-FEB-2000; 2000US-0496914.

-APR-2000; 2000US-0560875.

-JUN-2000; 2000US-0598075.

-JUL-2000; 2000US-0620325.

-SEP-2000; 2000US-0654936.

-SEP-2000; 2000US-0663561.

-OCT-2000; 2000US-0693325.

-NOV-2000; 2000US-0728422.

YSE-) HYSEQ INC.

I; 2001-476283/51.

PSDB; AAK52699.

cleic acids encoding polypeptides with cytokine-like activities,
 eful in diagnosis and gene therapy -

aim 20; Page 285-286; 6221pp; English.

e invention relates to polynucleotides (AAK51456-AAK53435) and the
 coded polypeptides (AAM78323-AAM80302) that exhibit activity elating to

CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.

XX SQ Sequence 654 AA;

Query Match	98.5%;	Score 2989.5;	DB 22;	Length 654;
Best Local Similarity	89.9%;	Pred. No. 6.9e-282;		
Matches 568;	Conservative 1;	Mismatches 0;	Indels 63;	Gaps 1;
QY	1	MDPAAVLQEKALFPM-----	-----	16
DB	23	MDPAAVLQEKALFPMFPRSCWCPGNTMAERSLTATSTSRVQCSMPRSLWLGCSLSADS	82	
QY	17	-----NSSREDNCNGEPPPKIIPKNSLRQTVNSCARLCLNQETV	57	
DB	83	MPSLACLNPCTGALTAFQNSSEREDNCNGEPPPKIIPKNSLRQTVNSCARLCLNQETV	142	
QY	58	CLASTAMKTENCVAKTCLANGTSSMIVPKQKLSASYEKEKELCVKYFEQWSESDDVEFY	117	
DB	143	CLASTAMKTENCVAKTCLANGTSSMIVPKQKLSASYEKEKELCVKYFEQWSESDDVEFY	202	
QY	118	EHLISOMCHYOHGHSNLYKPMLODFITLARGLDHIAENTLSYLDKSLCAAEVLCK	177	
DB	203	EHLISOMCHYOHGHSNLYKPMLODFITLARGLDHIAENTLSYLDKSLCAAEVLCK	262	
QY	178	EMRYVTSQGLMKWKKLIERMVTRTSLMRGLAERGGWQYLFKNKPPDGNAPPNSFYRALYP	237	
DB	263	EMRYVTSQGLMKWKKLIERMVTRTSLMRGLAERGGWQYLFKNKPPDGNAPPNSFYRALYP	322	
QY	238	KIIQDIETIESNRCGRHSLSQIHCRTSSTSGVYCLQYDDQKIVSGLRDNTIKIWDKNTL	297	
DB	323	KIIQDIETIESNRCGRHSLSQIHCRTSSTSGVYCLQYDDQKIVSGLRDNTIKIWDKNTL	382	
QY	298	ECKRIILTGTGTVLCLQYDERVLIITGSSDSTVRVWDVNTGEMLNTLIHHCCEAVLHLPFNN	357	
DB	383	ECKRIILTGTGTVLCLQYDERVLIITGSSDSTVRVWDVNTGEMLNTLIHHCCEAVLHLPFNN	442	
QY	358	GMMVTCSKDRSIAVMDASPTDITLRLVGVHRAAVNVVDFDDKIYIVSASGDRTIKWNT	417	
DB	443	GMMVTCSKDRSIAVMDASPTDITLRLVGVHRAAVNVVDFDDKIYIVSASGDRTIKWNT	502	
QY	418	STCEFVRLTNGHKGRIACLOVEDRLVSGSSDNTIRLWDIECGACLRVLEGHEELVRCIR	477	
DB	503	STCEFVRLTNGHKGRIACLOVEDRLVSGSSDNTIRLWDIECGACLRVLEGHEELVRCIR	562	
QY	478	FDNKRIVSGAYDGKIKVMDLVAALDPAPAGTLCRLTLVEHSGRVFRLQDFEFOIVSSSH	537	
DB	563	FDNKRIVSGAYDGKIKVMDLVAALDPAPAGTLCRLTLVEHSGRVFRLQDFEFOIVSSSH	622	
QY	538	DDTILWDFLNDPAAQAEPPSPSRTYIYSR	569	
DB	623	DDTILWDFLNDPAAQAEPPSPSRTYIYSR	654	

Search completed: October 22, 2003, 10:22:48

Job time : 86 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

tein - protein search, using sw model

: October 22, 2003, 10:25:22 ; Search time 72 Seconds
(without alignments)
1323.404 Million cell updates/sec

US-09-601-168b-2
t score: 3034
ce: 1 MDPAAVLQKALKFMSSE.....PAAQEPSPSRITYISR 569

g table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ed: 629382 seqs, 167460630 residues

number of hits satisfying chosen parameters: 629382

m DB seq length: 0

m DB seq length: 2000000000

rocessing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

se : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Score	Query Match	Length	DB ID	Description
3034	100.0	569	12	US-09-832-161-18
3034	100.0	569	14	US-10-042-417-2
3034	100.0	569	15	US-10-038-010-8
2384.5	78.6	542	12	US-09-832-161-16
1163	38.3	219	15	US-10-023-530-2
989.5	32.6	265	9	US-09-764-848-30
989.5	32.6	265	12	US-10-222-020-30
989.5	32.6	265	15	US-10-116-016-30
644.5	21.2	626	10	US-09-213-888-21
644.5	21.2	626	10	US-09-328-877A-21
640	21.1	540	10	US-09-213-888-7
640	21.1	540	10	US-09-213-888-10
640	21.1	540	10	US-09-328-877A-7
640	21.1	540	10	US-09-328-877A-10
640	21.1	540	12	US-10-245-618-14

Sequence 6, Appli
Sequence 6, Appli
Sequence 5, Appli
Sequence 8, Appli
Sequence 9, Appli
Sequence 8, Appli
Sequence 18, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 6, Appli
Sequence 27, Appli
Sequence 27, Appli
Sequence 25, Appli
Sequence 25, Appli
Sequence 2, Appli
Sequence 314, Appl
Sequence 3506, Ap
Sequence 8506, Ap
Sequence 30, Appl
Sequence 7658, Ap
Sequence 10, Appl
Sequence 29, Appl
Sequence 8, Appli
Sequence 12, Appli

ALIGNMENTS

RESULT 1
US-09-832-161-18
; Sequence 18, Application US/09832161
; Publication NO. US20030166587A1
; GENERAL INFORMATION:
; APPLICANT: Manning, Anthony M.
; APPLICANT: Mercutio, Frank
; APPLICANT: Amit, Sharon
; APPLICANT: Ben-Neriah, Yinon
; APPLICANT: Davis, Matti
; APPLICANT: Hatzubai, Ada
; APPLICANT: Lavon, Iris
; APPLICANT: Yaron, Avraham
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING ACTIVATION OF
; FILE REFERENCE: NP-KB
; CURRENT FILING DATE: 1998-12-10
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18

pcg date Sept 4 2003

pcg date Sept 4 2003

Query Match 100.0%; Score 3034; DB 12; Length 569;
Best Local Similarity 100.0%; Pred. No. 1.5e-290;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPAAVLQKALKFMSSEDCNNGEPKRIIPEKNSLRQTYNSCARLCLNQETVCLA 60
Db 1 MDPAAVLQKALKFMSSEDCNNGEPKRIIPEKNSLRQTYNSCARLCLNQETVCLA 60
Qy 61 STAMTENCVAKTLANGTSSMIVPKQKLSASKEKELCVKYPEQMSDQVEFVHL 120

61 STAMKTENCVAKTCLANGTSSMIVPKQKLSASYEKELCVKYFEQWSESDQVEFVEHL 120
121 ISQCHYQGHGHSYLSKPLMQDFITALPARGLDHIAENILSYLDAKSLCAELVCKEWY 180
121 ISQCHYQGHGHSYLSKPLMQDFITALPARGLDHIAENILSYLDAKSLCAELVCKEWY 180
181 RVTSDGMLWKLIJERMVTRDSLWRGLAERRGQGYLFKNKPPDGNAPPNSFYRALYPKII 240
181 RVTSDGMLWKLIJERMVTRDSLWRGLAERRGQGYLFKNKPPDGNAPPNSFYRALYPKII 240
241 ODITETESNWRGHSLSQRIHCRSETSGVYCLQYDDOKIVSGLRDNTIKIWDKNTLECK 300
241 ODITETESNWRGHSLSQRIHCRSETSGVYCLQYDDOKIVSGLRDNTIKIWDKNTLECK 300
301 RILTGTGSLVCLQYDERVERIITGSSSDSTVRVMDVNTGEMLNTLIHHCEAVLHFRFNGMM 360
301 RILTGTGSLVCLQYDERVERIITGSSSDSTVRVMDVNTGEMLNTLIHHCEAVLHFRFNGMM 360
361 VTCSKDRSIAVDMASPTDITLRRVLVGHRAAVNVDDDKYIIVSASGDRTIKWNNTSTC 420
361 VTCSKDRSIAVDMASPTDITLRRVLVGHRAAVNVDDDKYIIVSASGDRTIKWNNTSTC 420
421 EFVRTLNHKGRIACLOYRDRLVVGSSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN 480
421 EFVRTLNHKGRIACLOYRDRLVVGSSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN 480
481 KRIVSGAYDGKIKWDLVAALDPRAPAGTLCRLTLVHSGRVFRLOQDFEFOIVSSSHDDT 540
481 KRIVSGAYDGKIKWDLVAALDPRAPAGTLCRLTLVHSGRVFRLOQDFEFOIVSSSHDDT 540
541 ILIWDFLNDPAAQAEPSPRSRTYTIYSR 569
541 ILIWDFLNDPAAQAEPSPRSRTYTIYSR 569

2
042-417-2
ence 2, Application US/10042417
ication No. US20020123082A1
RAL INFORMATION:
LICANT: Pagano, M.
JE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF
JE OF INVENTION: PROLIFERATIVE AND DIFFERENTIATIVE DISORDERS
3 REFERENCE: 5914-030-999
ENT APPLICATION NUMBER: US/10/042,417
ENT FILING DATE: 2002-01-07
OR APPLICATION NUMBER: 60/260,179
OR FILING DATE: 2001-01-5
ER OF SEQ ID NOS: 89
TWARE: Patent In Ver. 2.0
ID NO 2
NGTH: 569
PE: PRT
GANISM: Homo sapiens
042-417-2

Match 100.0%; Score 3034; DB 14; Length 569;
Local Similarity 100.0%; Pred. No. 1.5e-290;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MDPAAVLQEKALKFNMSEREDCNGGPPRKIIPEKNSLRQTYNSCARLCLNOETVCLA 60
1 MDPAAVLQEKALKFNMSEREDCNGGPPRKIIPEKNSLRQTYNSCARLCLNOETVCLA 60
61 STAMKTENCVAKTCLANGTSSMIVPKQKLSASYEKELCVKYFEQWSESDQVEFVEHL 120
61 STAMKTENCVAKTCLANGTSSMIVPKQKLSASYEKELCVKYFEQWSESDQVEFVEHL 120
121 ISQCHYQGHGHSYLSKPLMQDFITALPARGLDHIAENILSYLDAKSLCAELVCKEWY 180
121 ISQCHYQGHGHSYLSKPLMQDFITALPARGLDHIAENILSYLDAKSLCAELVCKEWY 180
181 RVTSDGMLWKLIJERMVTRDSLWRGLAERRGQGYLFKNKPPDGNAPPNSFYRALYPKII 240

Db 181 RVTSDGMLWKLIJERMVTRDSLWRGLAERRGQGYLFKNKPPDGNAPPNSFYRALYPKII 240
Qy 241 ODITETESNWRGHSLSQRIHCRSETSGVYCLQYDDOKIVSGLRDNTIKIWDKNTLECK 300
Db 241 ODITETESNWRGHSLSQRIHCRSETSGVYCLQYDDOKIVSGLRDNTIKIWDKNTLECK 300
Qy 301 RILTGTGSLVCLQYDERVERIITGSSSDSTVRVMDVNTGEMLNTLIHHCEAVLHFRFNGMM 360
Db 301 RILTGTGSLVCLQYDERVERIITGSSSDSTVRVMDVNTGEMLNTLIHHCEAVLHFRFNGMM 360
Qy 361 VTCSKDRSIAVDMASPTDITLRRVLVGHRAAVNVDDDKYIIVSASGDRTIKWNNTSTC 420
Db 361 VTCSKDRSIAVDMASPTDITLRRVLVGHRAAVNVDDDKYIIVSASGDRTIKWNNTSTC 420
Qy 421 EFVRTLNHKGRIACLOYRDRLVVGSSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN 480
Db 421 EFVRTLNHKGRIACLOYRDRLVVGSSSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN 480
Qy 481 KRIVSGAYDGKIKWDLVAALDPRAPAGTLCRLTLVHSGRVFRLOQDFEFOIVSSSHDDT 540
Db 481 KRIVSGAYDGKIKWDLVAALDPRAPAGTLCRLTLVHSGRVFRLOQDFEFOIVSSSHDDT 540
Qy 541 ILIWDFLNDPAAQAEPSPRSRTYTIYSR 569
Db 541 ILIWDFLNDPAAQAEPSPRSRTYTIYSR 569

RESULT 3
US-10-038-010-8
; Sequence 8, Application US/10038010
; Publication No. US20030040089A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, Legrain
; TITLE OF INVENTION: Protein-protein interactions in adipocyte cells
; FILE REFERENCE: B4767A
; CURRENT APPLICATION NUMBER: US/10/038,010
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 60/259,377
; PRIOR FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 8
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: beta-TrCP1
; LOCATION: (1)..(569)
; OTHER INFORMATION:
US-10-038-010-8

Query Match 100.0%; Score 3034; DB 15; Length 569;
Best Local Similarity 100.0%; Pred. No. 1.5e-290;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MDPAAVLQEKALKFNMSEREDCNGGPPRKIIPEKNSLRQTYNSCARLCLNOETVCLA 60
Db 1 MDPAAVLQEKALKFNMSEREDCNGGPPRKIIPEKNSLRQTYNSCARLCLNOETVCLA 60
Qy 61 STAMKTENCVAKTCLANGTSSMIVPKQKLSASYEKELCVKYFEQWSESDQVEFVEHL 120
Db 61 STAMKTENCVAKTCLANGTSSMIVPKQKLSASYEKELCVKYFEQWSESDQVEFVEHL 120
Qy 121 ISQCHYQGHGHSYLSKPLMQDFITALPARGLDHIAENILSYLDAKSLCAELVCKEWY 180
Db 121 ISQCHYQGHGHSYLSKPLMQDFITALPARGLDHIAENILSYLDAKSLCAELVCKEWY 180
Qy 181 RVTSDGMLWKLIJERMVTRDSLWRGLAERRGQGYLFKNKPPDGNAPPNSFYRALYPKII 240
Db 181 RVTSDGMLWKLIJERMVTRDSLWRGLAERRGQGYLFKNKPPDGNAPPNSFYRALYPKII 240

241 QDIETIESNRCGRHSQRIHCRSETSGVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300
241 QDIETIESNRCGRHSQRIHCRSETSGVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300
301 RILTGTGTSVLCLOYDERVITIGSSDSTVRVMDVNTGEMMLTIHHCEAVLHLRFNNGM 360
301 RILTGTGTSVLCLOYDERVITIGSSDSTVRVMDVNTGEMMLTIHHCEAVLHLRFNNGM 360
361 VTCSKORSIAVMDMASPTDITLRVLVGHRAAVNVVDFDDKIVSASGDRITIKVWNTSTC 420
361 VTCSKORSIAVMDMASPTDITLRVLVGHRAAVNVVDFDDKIVSASGDRITIKVWNTSTC 420
421 EFVRTTLNGHKGRIACLOQRDLRVVSGSSDNTIRLWDIECGACLRVLEGGHEELVRCIRFDN 480
421 EFVRTTLNGHKGRIACLOQRDLRVVSGSSDNTIRLWDIECGACLRVLEGGHEELVRCIRFDN 480
481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTLVEHSGRVFRLQDFEQIVSSSHDDT 540
481 KRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTLVEHSGRVFRLQDFEQIVSSSHDDT 540
541 ILIWDFLNDPAAQAEPSPSRRTTYIISR 569
541 ILIWDFLNDPAAQAEPSPSRRTTYIISR 569

32-161-16
nce 16, Application US/09832161
cation No. US2003016587A1
AL INFORMATION:
ICANT: Manning, Anthony M.
ICANT: Mercurio, Frank
ICANT: Amit, Sharon
ICANT: Ben-Neriah, Yinon
ICANT: Davis, Metti
ICANT: Hatzubai, Ada
ICANT: Lavon, Iris
ICANT: Yaron, Avraham
E OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING ACTIVATION OF
E OF INVENTION: NF-kB
REFERENCE: 860098.427
ENT APPLICATION NUMBER: US/09/832,161
ENT FILING DATE: 2001-04-09
2 APPLICATION NUMBER: 09/210,060
2 FILING DATE: 1998-12-10
ER OF SEQ ID NOS: 30
ARE: Patent in Ver. 2.0
3 NO 16
3TH: 542
5: PRT
ANISM: Homo sapiens
32-161-16

Match 78.6%; Score 2384.5; DB 12; Length 542;
Local Similarity 79.0%; Pred. No. 2e-226;
as 451; Conservative 48; Mismatches 41; Indels 31; Gaps 7;
1 MDPAEAVLQKALKFNNSSEREDCNNGBPBKIIPEKNSLRQTYNSCARCLNQETVCLA 60
1 MEP-DSVIEDKTIELMCS-----VP-----RSLWGLCANLV---ESMCAL 36
61 S--TAMKTENCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKYPEOWSESQVDFVEHL 118
37 SCLQSPSVKCL---QISNQTSSVIVSRKRPSSGNTQKENDLCIKYFDQWSESQVDFVE 93
119 HLISQCHYQHGHINSYLPKMLQDFITLAPGLDHTAENILSYLDAKSLCAAEVLCKE 178
94 HLISRMCHYQHGHINSYLPKMLQDFITLAPGLDHTAENILSYLDAKSLCAAEVLCKE 153
179 WYRVTSQGLMKWLKLIERNVTRDSLWGLAERRGQVLPKPKPDGNAPNPFYALYPK 238
154 WQVVISRQGLMKWLKLIERNVTRDSLWGLAERRGQVLPKPKRPTDG--PPNSFYRSLYPK 211

Qy 239 IIODIETIESNRCGRHSQRIHCRSETSGVYCLQYDDQKIVSGLRDNTIKIWDKNTLE 298
Db 212 IIODIETIESNRCGRHNLRQCRSENSKGVYCLQYDEKIIISGLRNSIKIWDKTSLE 271
Qy 299 CKRILTGTGTSVLCLOYDERVITIGSSDSTVRVMDVNTGEMMLTIHHCEAVLHLRFNNG 358
Db 272 CLKVLGTGTSVLCLOYDERVITIGSSDSTVRVMDVNTGEMMLTIHHCEAVLHLRFNSG 331
Qy 359 MMVTCSDRSIAVMDMASPTDITLRVLVGHRAAVNVVDFDDKIVSASGDRITIKVWNTS 418
Db 332 LMVTCSDRSIAVMDMASPTDITLRVLVGHRAAVNVVDFDDKIVSASGDRITIKVWNTS 391
Qy 419 TCFVRTTLNGHKGRIACLOQRDLRVVSGSSDNTIRLWDIECGACLRVLEGGHEELVRCIRF 478
Db 392 TCFVRTTLNGHKGRIACLOQRDLRVVSGSSDNTIRLWDIECGACLRVLEGGHEELVRCIRF 451
Qy 479 DNKRIVSGAYDGKIKVMDLVAALDPRAPAGTLCRLTLVEHSGRVFRLQDFEQIVSSSHD 538
Db 452 DNKRIVSGAYDGKIKVMDLQAALDPRAPASTLCRLTLVEHSGRVFRLQDFEQIISSSH 511
Qy 539 DTILWDFLNDPAAQAEPSPSRRTTYIISR 569
Db 512 DTILWDFLNPSPAQNETRSPSRRTTYIISR 542

RESULT 5
US-10-023-530-2
; Sequence 2, Application US/10023530
; Publication No. US20030007956A1
; GENERAL INFORMATION:
; APPLICANT: LEGRAIN, Pierre
; APPLICANT: BENAROUS, Richard
; APPLICANT: BLOT, Guillaume
; APPLICANT: LASSOT, Irina
; TITLE OF INVENTION: PROTEINS THAT INTERACT WITH BETA TRCP
; FILE REFERENCE: B4717A
; CURRENT APPLICATION NUMBER: US/10/023,530
; CURRENT FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: 60/256,276
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: BetaTrCP
; LOCATION: (1)..(219)
; OTHER INFORMATION: P-box protein
US-10-023-530-2

Query Match 38.3%; Score 1163; DB 15; Length 219;
Best Local Similarity 100.0%; Pred. No. 2.2e-106;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MDPAEAVLQKALKFNNSSEREDCNNGBPBKIIPEKNSLRQTYNSCARCLNQETVCLA 60
Db 1 MDPAEAVLQKALKFNNSSEREDCNNGBPBKIIPEKNSLRQTYNSCARCLNQETVCLA 60
Qy 61 STAMKTENCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKYPEOWSESQVDFVEHL 120
Db 61 STAMKTENCVAKTKLANGTSSMIVPKORKLSASYEKEKELCVKYPEOWSESQVDFVEHL 120
Qy 121 ISQWCHYQHGHINSYLPKMLQDFITLAPGLDHTAENILSYLDAKSLCAAEVLCKEY 180
Db 121 ISQWCHYQHGHINSYLPKMLQDFITLAPGLDHTAENILSYLDAKSLCAAEVLCKEY 180
Qy 181 RVTSQGLMKWLKLIERNVTRDSLWGLAERRGQVLPK 219
Db 181 RVTSQGLMKWLKLIERNVTRDSLWGLAERRGQVLPK 219

[illegible]

DR APPLICATION NUMBER: 60/236,370
DR FILING DATE: 2000-09-29
DR APPLICATION NUMBER: 60/236,802
DR FILING DATE: 2000-10-02
DR APPLICATION NUMBER: 60/237,037
DR FILING DATE: 2000-10-02
DR APPLICATION NUMBER: 60/237,040
DR FILING DATE: 2000-10-02
DR APPLICATION NUMBER: 60/240,960
DR FILING DATE: 2000-10-20
DR APPLICATION NUMBER: 60/239,935
DR FILING DATE: 2000-10-13
DR APPLICATION NUMBER: 60/239,937
DR FILING DATE: 2000-10-13
DR APPLICATION NUMBER: 60/241,787
DR FILING DATE: 2000-10-20
DR APPLICATION NUMBER: 60/246,474
DR FILING DATE: 2000-11-08
DR APPLICATION NUMBER: 60/246,532
DR FILING DATE: 2000-11-08
DR APPLICATION NUMBER: 60/249,216
DR FILING DATE: 2000-11-17
DR APPLICATION NUMBER: 60/249,210
DR FILING DATE: 2000-11-17
DR APPLICATION NUMBER: 60/226,681
DR FILING DATE: 2000-08-22
DR APPLICATION NUMBER: 60/225,759
DR FILING DATE: 2000-08-14
DR APPLICATION NUMBER: 60/225,213
DR FILING DATE: 2000-08-14
DR APPLICATION NUMBER: 60/227,182
DR FILING DATE: 2000-08-22
DR APPLICATION NUMBER: 60/225,214
DR FILING DATE: 2000-08-14
DR APPLICATION NUMBER: 60/235,836
DR FILING DATE: 2000-09-27
R APPLICATION NUMBER: 60/230,438
R FILING DATE: 2000-09-06
R APPLICATION NUMBER: 60/215,135
R FILING DATE: 2000-06-30
R APPLICATION NUMBER: 60/225,266
R FILING DATE: 2000-08-14
R APPLICATION NUMBER: 60/249,218
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,208
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,213
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,212
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,207
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,245
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,244
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,217
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,211
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,215
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,264
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,214
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/249,297
R FILING DATE: 2000-11-17
R APPLICATION NUMBER: 60/232,400
R FILING DATE: 2000-09-14
R APPLICATION NUMBER: 60/231,242
R FILING DATE: 2000-09-08
R APPLICATION NUMBER: 60/232,081

DR FILING DATE: 2000-09-08
DR APPLICATION NUMBER: 60/232,080
DR FILING DATE: 2000-09-08
DR APPLICATION NUMBER: 60/231,414
DR FILING DATE: 2000-09-08
DR APPLICATION NUMBER: 60/231,244
DR FILING DATE: 2000-09-08
DR APPLICATION NUMBER: 60/233,064
DR FILING DATE: 2000-09-14
DR APPLICATION NUMBER: 60/233,063
DR FILING DATE: 2000-09-14
DR APPLICATION NUMBER: 60/232,397
DR FILING DATE: 2000-09-14
DR APPLICATION NUMBER: 60/232,399
DR FILING DATE: 2000-09-14
DR APPLICATION NUMBER: 60/232,401
DR FILING DATE: 2000-09-14
DR APPLICATION NUMBER: 60/241,808
DR FILING DATE: 2000-10-20
DR APPLICATION NUMBER: 60/241,826
DR FILING DATE: 2000-10-20
DR APPLICATION NUMBER: 60/241,786
DR FILING DATE: 2000-10-20
DR APPLICATION NUMBER: 60/241,221
DR FILING DATE: 2000-10-20
DR APPLICATION NUMBER: 60/246,475
DR FILING DATE: 2000-11-08
Query Match 32.6%; Score 989.5; DB 12; Length 265;
Best Local Similarity 68.9%; Pred. No. 4.1e-89;
Matches 186; Conservative 24; Mismatches 21; Indels 39; Gaps 3;
QY 15 FMNSEREDCNNGEPKRIIPEKNSLQTYNSCARLCLNQETVCLASTAMKTENCVAKTK 74
DB 20 FQNTSVMEDQNEDESPK-----KNTLWQ----- 42
QY 75 LANGTSSMIVPKQKLSASVEKEKELCVKFEQWSESDQVEFVEHLISQWCHYQHGHNS 134
DB 43 ISNGTSSVIVSRKRPSEGNQKQKDLCKIFYDQWSESDQVEFVEHLISRMCHYQHGHNS 102
QY 135 YLKPMLQDFITALPARGLDHIAENILSYLDAKSLCAELVCKEWYRVTSQGLMKWKLIE 194
DB 103 YLKPMLQDFITALPEQGLDHIENILSYLDAKSLCAELVCKEWQVISEGMLWKLIE 162
QY 195 RMVRTDSLWRGLAERCGWQYLFKNKPPDGNAPNSFYRALYFKPIQDIETIESNWRCCR 254
DB 163 RMVRTDPLWKGLSERGWDQYLFKNRPTDG--PPNSFYRSLYKPIQDIETIESNWRCCR 220
QY 255 HSLQRIHCRSETSKGVYCLQYDQKIVSGL 284
DB 221 HNLQRIQCRSENSKGVYCLQYDDEKIISGL 250
RESULT 8
US-10-116-016-30
Sequence 30, Application US/10116016
Publication No. US20030054379A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PTZ08C1
CURRENT APPLICATION NUMBER: US/10/116,016
CURRENT FILING DATE: 2002-04-05
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 53
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 30
LENGTH: 265
TYPE: PRI
ORGANISM: Homo sapiens
US-10-116-016-30
Query Match 32.6%; Score 989.5; DB 15; Length 265;

```
Local Similarity 68.9%; Pred. No. 4.1e-89;
hes 186; Conservative 24; Mismatches 21; Indels 39; Gaps 3;
15 PMNSEREDCNNGEPKPIPEKSLRQTYNSCARCLNQETVCLASTAMKTENCYAKTK 74
20 PONTSMEDQNEDESPK-----KNTLWQ-----42
75 LANGTSSMIVPKQKLSASYEKELCVKYPEQWSESDQVEFVEHLISQCHYQGHINS 134
43 ISNGTSSVIVSRKPSSEGNCKEDLCIKYFDQWSESDQVEFVEHLISRCHYQGHINS 102
135 YLKPMQLQDFITALPARGLDHIAENILSYLDAKSLCAELVCWEYRYVTSQGLMKXLE 194
103 YLKPMQLQDFITALPEQGLDHAENILSYLDAKSLCAELVCWEYRYVTSQGLMKXLE 162
195 RMVRTDSLRGLAERRGQYLFKNKPPDGNAPNSFVRALYPKIIODIETIESNWRCCR 254
163 RMVRTDPLWKLSERRGWQYLFKNRPTDG--PNSFYSRLYPKIIODIETIESNWRCCR 220
255 HSLQRIHCRSETSGVYCLQYDDQKIVSGL 284
221 HNLQRIQCRSENSGVYCLQYDDQKIVSGL 250
9
213-888-21
nce 21, Application US/09213888A
nt No. US20020164683A1
AL INFORMATION:
JICANT: Gurney, Mark E.
JICANT: Li, Jinhe
JICANT: Pauley, Adele M.
JICANT: Pharmacia & Upjohn Company
E OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
E OF INVENTION: Encode Them
3 REFERENCE: 6142
RENT APPLICATION NUMBER: US/09/213.888A
RENT FILING DATE: 1998-12-17
ER OF SEQ ID NOS: 27
WARE: PatentIn Ver. 2.0
D NO 21
GTH: 626
E: PRT
ANISM: Artificial Sequence
TURE:
IER INFORMATION: Description of Artificial Sequence: 6 myc tagged
IER INFORMATION: homo sapien
213-888-21
Match
Local Similarity 21.2%; Score 644.5; DB 10; Length 626;
hes 169; Conservative 103; Mismatches 22; Indels 67; Gaps 15;
8 LOEKALKFPMNSEREDCNNGEPKPIPEK--NSLRQ--TYNSCARCLNQETVCLASTA 63
53 MEQKLISEEDLNEMESLGLTMEQKLISEEDLNSMKRKLHDGSEVRSFSLGKPKVSEY 112
64 MXTENCV-----AKTKLANGTSSMIVPKQKLSASYEKEL--CVKYPEQWSESD 172
113 TSTTGLVPCSATPTTFGLDRAANGQ-----QQRRTTSVQPTGLQEWLKMFSQSGPE 167
113 QVEFVEHLISQCHYQGHINSYLPKMLQDFITALPARGLDHIAENILSYLDAKSLCAA 172
168 KLLALDELIDSCPTQVKHMQVIEPQFQDFISLIP-----KELALYVLSFLEPKDLQA 223
173 ELVCKEWTRVTSQGLMKXLERMVVRTDSLRGLAE-----RRGWGYLFKNKPPDGNAP 227
224 AQTCEYRWILAEDNLLWRECKE-----EGIDEPLHIKRRK-----VIKPGFIHSP 269
228 PMSFVRALYPKIIODIETIESNWRCCRHSRLQRIHCRSETSGVYCLQYDDQKIVSGLRD 287
270 WKSAY-----IRQ--HRIDTNWRGELKSPKV-LKGHDDHVIITCLOFCGNRIVSGSDDN 320
288 TIKINDKNTLECKRLITLGTGTVCLQYDERVITGSSDSTVRVMDVNTGEMLNTLIHHC 347
321 TLKWSAVTGKCLRTLVTGHTGGVSSQMRDNIISGSDTRTLKVNNAETGECIHTLYGHT 380
```

348 EAVLHLRNNMMWTCSDKRSIAVWDMASPTDITLRLVVGHRRAAVNVVDFDDKYIVSAS 407
381 STVRCMELHKKRVVSSGRDATLRWDIETGQCL---HVLGMHVAARVCQYDGRVYVSGA 437
408 GDRITKWNSTCEFTVLNKHKGIAQYRORLVVSGSSDNTIRLWIECCACLRVLE 467
438 YDFNVKWDPEFETCLTLOGHTRVVSLOFDGHHVSSGLDTSIRVWDVETGNCIHTLT 497
468 GHEELVRCIRPDNKRIVSGYDGGIKVWDLVAALDPRAPAGTLCRLTLV---EHSGRVFR 524
498 GHOSLTSGMELKDNILVSGNADSVTKIWDIKTG-----QCLQTLOGPNKHQSANTC 548
525 LQDFEQIVSSSHDDTLIWD 545
549 LQFNKNFVITSSDDGTGVLMD 569

213-888-7
ance 7, Application US/09213888A
ic No. US20020164683A1
RAL INFORMATION:
ICANT: Gurney, Mark E.
ICANT: Li, Jinhe
ICANT: Pauley, Adele M.
ICANT: Pharmacia & Upjohn Company
E OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
E OF INVENTION: Encode them
REFERENCE: 6142
ENT APPLICATION NUMBER: US/09/213.888A
ENT FILING DATE: 1998-12-17
ER OF SEQ ID NOS: 27
WARE: Patent In Ver. 2.0
D NO 7
GTH: 540
E: PRT
ANISM: Homo sapiens
13-888-7

Match 21.1%; Score 640; DB 10; Length 540;
Local Similarity 30.8%; Pred. No. 4.1e-54;
es 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;
14 KFMSSREDNCNGEPKRIPEKNLSRQTYNSCARLCLNQETVCLASTAMKTCNCVAKT 73
4 KLDHGEVRSFSLGKKCKV-----SEYTTTGL-----VPCSA-----TPTTFGDL 45
74 KLANGTSSMIVPKORKLSASVEKEKEL---CVKYFEQMSQSDQVEFVHLISQMHYQGH 131
46 RAANGQG-----QQRRTITSVQPTGLQEWLKFQSWGPEKLLALDELIDSCPTQVKH 100
132 INSVLKPMLODFITALPARGLDHIAENILSYLDAKSLCAELVCKEWYVTS DGLMVK 191
101 MMQVIEPQDFDLSLLP-----KELALYVLSFLEPKDLQAQTCRYWRILAEDNLLWRE 156
192 LIERMVRTDSLWRGLAE-----RRGWQYLFKNKPPDGNAPPNSFYRALYPKIIQDIETI 246
157 KCKE-----EGIDEPLHIKRRK-----VIKPGFIHSPWKSAY-----IRQ---HRI 194
247 ESNWRCGRHSIQRHCRSETSGYVCIQYDDOKIVSGLRNTIKIWNKNTLECKRLITGH 306
195 DTNWRRGELKSPKV-LKGHDHVTICLOFCGNRIVSGSDNTLTKVWSAVTGKCLRTLUGH 253
307 TGSVLCLQYDREVIITGSSDSTVRVWDVNTGEMLNTLIHCEAVLHLRFNNGMMVTCSD 366
254 TGGVWSSQWRNIIISGSDTLKVMNAETGECIHTLYGHTSTVRCMHLHKKRVVSGSRD 313
367 RSIADVMDASPTDITLRLVVGHRRAAVNVVDFDDKYIVSASGDRITIKWNNTSTCBFVRTL 426
314 ATLRWMDIETGQCL---HVLGMHVAARVCQYDGRVYVSGAYDPMVKVMDPETETCLATL 370
427 NGHKEGIAQYRDELVYVSGSSDNTIRLWIECCACLRVLEGEELVRCIRFNDKRIVSG 486

Db 371 QGHTNRVYSLOFDGHHVSSGLDTSIRVWDVETGNCIHTLTGHOSLTSGMELKDNILVSG 430
Qy 487 AYDGKIKVWDLVAALDPRAPAGTLCRLTLV---EHSGRVFRLOFDEFQIVSSSHDDTLI 543
Db 431 NADSTVKIWDIKTG-----QCLQTLOGPNKHQSANTCLOFNKNFVITSSDDGTGVLK 481
Qy 544 WD 545
Db 482 WD 483

RESULT 12
US-09-213-888-10
; Sequence 10, Application US/09213888A
; Patent No. US20020164683A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/213.888A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 10
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-213-888-10

Query Match 21.1%; Score 640; DB 10; Length 540;
Best Local Similarity 30.8%; Pred. No. 4.1e-54;
Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;
Qy 14 KFMSSREDNCNGEPKRIPEKNLSRQTYNSCARLCLNQETVCLASTAMKTCNCVAKT 73
Db 4 KLDHGEVRSFSLGKKCKV-----SEYTTTGL-----VPCSA-----TPTTFGDL 45
Qy 74 KLANGTSSMIVPKORKLSASVEKEKEL---CVKYFEQMSQSDQVEFVHLISQMHYQGH 131
Db 46 RAANGQG-----QQRRTITSVQPTGLQEWLKFQSWGPEKLLALDELIDSCPTQVKH 100
Qy 132 INSVLKPMLODFITALPARGLDHIAENILSYLDAKSLCAELVCKEWYVTS DGLMVK 191
Db 101 MMQVIEPQDFDLSLLP-----KELALYVLSFLEPKDLQAQTCRYWRILAEDNLLWRE 156
Qy 192 LIERMVRTDSLWRGLAE-----RRGWQYLFKNKPPDGNAPPNSFYRALYPKIIQDIETI 246
Db 157 KCKE-----EGIDEPLHIKRRK-----VIKPGFIHSPWKSAY-----IRQ---HRI 194
Qy 247 ESNWRCGRHSIQRHCRSETSGYVCIQYDDOKIVSGLRNTIKIWNKNTLECKRLITGH 306
Db 195 DTNWRRGELKSPKV-LKGHDHVTICLOFCGNRIVSGSDNTLTKVWSAVTGKCLRTLUGH 253
Qy 307 TGSVLCLQYDREVIITGSSDSTVRVWDVNTGEMLNTLIHCEAVLHLRFNNGMMVTCSD 366
Db 254 TGGVWSSQWRNIIISGSDTLKVMNAETGECIHTLYGHTSTVRCMHLHKKRVVSGSRD 313
Qy 367 RSIADVMDASPTDITLRLVVGHRRAAVNVVDFDDKYIVSASGDRITIKWNNTSTCBFVRTL 426
Db 314 ATLRWMDIETGQCL---HVLGMHVAARVCQYDGRVYVSGAYDPMVKVMDPETETCLHTL 370
Qy 427 NGHKEGIAQYRDELVYVSGSSDNTIRLWIECCACLRVLEGEELVRCIRFNDKRIVSG 486
Db 371 QGHTNRVYSLOFDGHHVSSGLDTSIRVWDVETGNCIHTLTGHOSLTSGMELKDNILVSG 430
Qy 487 AYDGKIKVWDLVAALDPRAPAGTLCRLTLV---EHSGRVFRLOFDEFQIVSSSHDDTLI 543
Db 431 NADSTVKIWDIKTG-----QCLQTLOGPNKHQSANTCLOFNKNFVITSSDDGTGVLK 481

```
544 WD 545
||
482 WD 483

328-877A-7
ence 7, Application US/09328877A
nt No. US20020177187A1
RAL INFORMATION:
LICANT: Gurney, Mark E.
LICANT: Li, Jinhe
LICANT: Pauley, Adele M.
LICANT: Pharmacia & Upjohn Company
E OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
E OF INVENTION: Encode Them
REFERENCE: 6142
CURRENT APPLICATION NUMBER: US/09/328, 877A
CURRENT FILING DATE: 1999-06-09
NUMBER OF SEQ ID NOS: 27
WARE: PatentIn Ver. 2.0
ID NO 7
GTH: 540
E: PRT
ANISM: Homo sapiens
328-877A-7

/ Match 21.1%; Score 640; DB 10; Length 540;
Local Similarity 30.8%; Pred. No. 4.1e-54;
Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

14 KFMNSREDNCNGEPPRKIIPEKNSLRQTYNSCARLCLNQETVCLASTAMKTCVAKT 73
4 KLDHGSSEVRSPSLGKKPKV-----SEYTTTGL-----VPCSA-----TPTTFGDL 45

74 KLANGTSSMIVPORKLSASYEKEKEL--CVKYFEQWSESQDOVEFVEHLISOMCHYQGHG 131
46 RAANQG-----QRRRTSVQPTGLQEWLKFQSWGPEKLLALDELIDSCPTQVRK 100

132 INSVLKPMLODFITALPARGLDHIAENILSYLDAKSLCAAEVLCKEYRVTSDGMLWKK 191
101 MMQVIEPQFQDFISLLP-----KELALYVLSFLEPKDLLQAAQTCRYWRLAEDNLLWRE 156
192 LIERNVRTDSLWRGLAE-----RRGWGYLFKNKPPDGNAPPNSFYRALYPKIIQDIETI 246
157 KCKE-----EGIDEPLHIKRRK-----VIKPGFIHSPKSA-----IRQ--HRI 194
247 ESNWRCGRHSIQRHCRSETSGKYVCLQYDDQKIVSGLRDNTIKIWDKNTLECKRIITGH 306
195 DTNWRGELKSPKV-LKGHDDHVITCLQFCGNRIVSGSDNTLVKWSAVTGKCLRTLAVGH 253
307 TGSVLCLOYDERVLIITGSSDSTVRVWDVNTGEMLNTLIHCEAVLHLRFNNGMMVTCSD 366
254 TGGWSSQMRDNIISGSTDRTLKWNNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRD 313
367 RSIADVMDASPTDITLRLVVGHRAAVNVVDFDKYIVSASGDRTIKVMNTSTCEFVRTL 426
314 ATLAVWDIETGQCL---HVLGMHVAARVCYQDGRVVSAYDFMVKVMWDPETETCLHTL 370
427 NGHKRGIAQLQYRDLRVVSGSDNTIRLWDIECGACLRVLEGEHELVRCTRFNKRIVSG 486
371 QGHTNRVYSLOFDGIHVVSGSLDTSIRVMDVETGNCIHTLTGHSLTSGMELKDNILVSG 430
487 AYDGKIKVMDLVAALDPRAPAGTLCRLTLV---EHSGRVPRLOQDFEFOIVSSSHDDTILI 543
431 NADSVTKIWDIKTG-----OCLQTLQGNKHQSVAVTCLOFNKNFVITSSDDGTVKL 481

544 WD 545
||
482 WD 483

US-09-328-877A-10
; Sequence 10, Application US/09328877A
; Patent No. US20020177187A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; TITLE OF INVENTION: Encode Them
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/328, 877A
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-328-877A-10

Query Match 21.1%; Score 640; DB 10; Length 540;
Best Local Similarity 30.8%; Pred. No. 4.1e-54;
Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

QY 14 KFMNSREDNCNGEPPRKIIPEKNSLRQTYNSCARLCLNQETVCLASTAMKTCVAKT 73
DB 4 KLDHGSSEVRSPSLGKKPKV-----SEYTTTGL-----VPCSA-----TPTTFGDL 45
QY 74 KLANGTSSMIVPORKLSASYEKEKEL--CVKYFEQWSESQDOVEFVEHLISOMCHYQGHG 131
DB 46 RAANQG-----QRRRTSVQPTGLQEWLKFQSWGPEKLLALDELIDSCPTQVRK 100
QY 132 INSVLKPMLODFITALPARGLDHIAENILSYLDAKSLCAAEVLCKEYRVTSDGMLWKK 191
DB 101 MMQVIEPQFQDFISLLP-----KELALYVLSFLEPKDLLQAAQTCRYWRLAEDNLLWRE 156
QY 192 LIERNVRTDSLWRGLAE-----RRGWGYLFKNKPPDGNAPPNSFYRALYPKIIQDIETI 246
DB 157 KCKE-----EGIDEPLHIKRRK-----VIKPGFIHSPKSA-----IRQ--HRI 194
QY 247 ESNWRCGRHSIQRHCRSETSGKYVCLQYDDQKIVSGLRDNTIKIWDKNTLECKRIITGH 306
DB 195 DTNWRGELKSPKV-LKGHDDHVITCLQFCGNRIVSGSDNTLVKWSAVTGKCLRTLAVGH 253
QY 307 TGSVLCLOYDERVLIITGSSDSTVRVWDVNTGEMLNTLIHCEAVLHLRFNNGMMVTCSD 366
DB 254 TGGWSSQMRDNIISGSTDRTLKWNNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRD 313
QY 367 RSIADVMDASPTDITLRLVVGHRAAVNVVDFDKYIVSASGDRTIKVMNTSTCEFVRTL 426
DB 314 ATLAVWDIETGQCL---HVLGMHVAARVCYQDGRVVSAYDFMVKVMWDPETETCLHTL 370
QY 427 NGHKRGIAQLQYRDLRVVSGSDNTIRLWDIECGACLRVLEGEHELVRCTRFNKRIVSG 486
DB 371 QGHTNRVYSLOFDGIHVVSGSLDTSIRVMDVETGNCIHTLTGHSLTSGMELKDNILVSG 430
QY 487 AYDGKIKVMDLVAALDPRAPAGTLCRLTLV---EHSGRVPRLOQDFEFOIVSSSHDDTILI 543
DB 431 NADSVTKIWDIKTG-----OCLQTLQGNKHQSVAVTCLOFNKNFVITSSDDGTVKL 481
QY 544 WD 545
DB 482 WD 483

RESULT 15
US-10-245-618-14
; Sequence 14, Application US/10245618
; Publication No. US20030143582A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven
```

LICANT: Strohmaier, Heimo
LICANT: Spruck, Charles
LICANT: Sangfelt, Olie
UE OF INVENTION: HCDC4 MODULATES CYCLIN E DEGRADATION
E REFERENCE: TSRI 779.2
RENT APPLICATION NUMBER: US/10/245,618
RENT FILING DATE: 2002-09-16
OR APPLICATION NUMBER: US 60/404,116
OR FILING DATE: 2002-08-15
OR APPLICATION NUMBER: US 60/322,947
OR FILING DATE: 2001-09-14
BER OF SEQ ID NOS: 48
WARE: FastSeq for Windows Version 4.0
ID NO 14
NGTH: 540
PE: PRT
GANISM: Artificial Sequence
ATURE:
HER INFORMATION: Synthesized
245-618-14
/ Match 21.1%; Score 640; DB 12; Length 540;
Local Similarity 30.8%; Pred. No. 4.1e-54;
es 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;
14 KPMSSEREDCNGEPPEPKIIPEKNSLRQTYNSCARCLNQETVCLASTAMKTCNCVAKT 73
4 KLDHGSEVRSFSLGKKPKCV-----SEYTSITGL-----VPCSA-----TPTFGDL 45
74 KLANGTSMIVPKQKLSASVEKSEL--CVKYFEQMSSESDQVEFVEHLISQMCVYQHG 131
46 RAANGCG-----QQRRTTSVQPTTGLQEWLKNFTQSWSGPEKLLALDELIDSCPTGVKH 100
132 INSVLKPMLQRFITALPARGLDHAEINILSVLDAKSICAAELVCKEWRYVTSQMLWKK 191
101 MNQVIEPQOFQDFISLLP-----KELALVLSFLBPKDLQAAQTCRYWRILAEONLLWRE 156
192 LIERMVRTDSLWRGLAE-----RRWGQYLFKNKFPDGNAPPNSFYALYKPKIIODIETI 246
157 KCKE-----EGIDEPLHIKRRK-----VIKPGFTHSPKSAV-----IRQ--HRI 194
247 ESNWRCGRHSIORIHCRSETSGVYCYQYDDQKIVSGLRDNITIKWDKNTLECKRLTGH 306
195 DTNWRGELKSPKV-LKGDDHVIITCQFCGNRIVSGDDNTLRKWSAVTGKCLRTLUGH 253
307 TGSVLCQYDERVIITGSSDSTVRVWDYNTGEMLNTLIHCEAVLHLRFNNGMMVTCSD 366
254 TGVWSSQWRDNIISGSTDRTLKWNAAETGECIHTLYGHTSTVRCMELHEKRVVSGSRD 313
367 RSIAYWDMASPTDITLRVLYCHRAAVNVDPDDKYIVSASGDRTIKWNTSICEFYRTL 426
314 ATRLVWDIETGQCL---HVLGHWAAVRCVQYDGRVVRVSGAYDFWVKVDPETETCUHTL 370
427 NGHKRGIAQLQYRDLVYSGSDNTIRLWDIECGACLRVLEGGHEELVRCIRFONKRVISG 486
371 QGHTNRVYSLQPDGHIHVVSGLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSG 430
487 AYDGKIKWDLVAALDPRAPACTLCRLTV---EHSGRVFLQDFEQIVSSSHDDTILI 543
431 NADSTVKIWDIKTG-----OCLQTLQGNPKHQSAVTCLOFNKNFVITSSDDGTVKL 481
544 WD 545
482 WD 483

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

tein - protein search, using sw model

October 22, 2003, 10:21:24 ; Search time 30 Seconds
(without alignments)
802.496 Million cell updates/sec

us-09-601-168b-2
score: 3034
1 MDPAAVLQEKALKFNSSP.....PAAQAPPPSPRTVTYISR 569
BLOSUM62
Gapop 10.0 , Gapext 0.5
328717 seqs, 42310858 residues

number of hits satisfying chosen parameters: 328717

DB seq length: 0

DB seq length: 2000000000

rocessing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pcp.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pcp.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pcp.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pcp.*
5: /cgn2_6/ptodata/2/iaa/PTUS COMB.pcp.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pcp.*

red. No. is the number of results predicted by chance to have a
core greater than or equal to the score of the result being printed,
nd is derived by analysis of the total score distribution.

SUMMARIES

Score	Query Match	Length	DB ID	Description
2582.5	85.1	517	1	US-08-190-802A-30
2582.5	85.1	517	3	US-08-477-346-30
2582.5	85.1	517	4	US-08-473-089-30
2582.5	85.1	517	4	US-08-487-072A-30
545	18.0	640	4	US-08-177-165A-30
520	17.1	587	3	US-08-899-578-2
399	13.2	779	1	US-08-190-802A-32
399	13.2	779	3	US-08-477-346-32
399	13.2	779	4	US-08-473-089-32
399	13.2	779	4	US-08-487-072A-32
399	13.2	779	4	US-08-177-165A-29
392	12.9	732	4	US-08-914-999-8
354	11.7	409	2	US-08-283-917-3
354	11.7	409	2	US-08-961-716-3
354	11.7	410	2	US-08-283-917-9
354	11.7	410	2	US-08-961-716-9
339.5	11.2	409	1	US-08-190-802A-51
339.5	11.2	409	3	US-08-477-346-51
339.5	11.2	409	4	US-08-473-089-51
339.5	11.2	409	4	US-08-487-072A-51
337	11.1	1146	4	US-08-914-999-6
321.5	10.6	514	1	US-08-190-802A-66
321.5	10.6	514	3	US-08-477-346-66
321.5	10.6	514	4	US-08-473-089-66
321.5	10.6	514	4	US-08-487-072A-66
318	10.5	422	1	US-08-190-802A-52
318	10.5	422	3	US-08-477-346-52

Sequence 52, Appl
Sequence 52, Appl
Sequence 6, Appl
Sequence 5, Appl
Sequence 5, Appl
Sequence 62, Appl
Sequence 3, Appl
Sequence 62, Appl
Sequence 62, Appl
Sequence 62, Appl
Sequence 2, Appl
Sequence 26, Appl
Sequence 2, Appl
Sequence 16, Appl
Sequence 16, Appl
Sequence 18, Appl
Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-08-130-802A-30
; Sequence 30, Application US/08190802A
; Patent No. 5519003
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: MD-40 - Derived Peptides and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-0850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,802A
; FILING DATE: 01-FEB-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 8600-0139
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 517 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13
US-08-190-802A-30

Query Match 85.1%; Score 2582.5; DB 1; Length 517;
Best Local Similarity 91.4%; Pred. No. 3.le-266;
Matches 487; Conservative 7; Mismatches 8; Indels 31; Gaps 2;
18 SSERDCNNGEPFKIPEKNSLRQTYNSCARLCLNQETVCLASTAMKTCNVAKTKLAN 77

SOFTWARE: PatentIn Release #1.0, Version #1.25

RENT APPLICATION DATA:

APPLICATION NUMBER: US/08/473,089

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

TORNEY/AGENT INFORMATION:

NAME: MURASHIGE, KATE H.

REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 2550-0025.22

TELEPHONE: (202) 887-1500

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/473,072A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: MURASHIGE, KATE H.

REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 2550-0025.20

TELEPHONE: (202) 887-1500

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500

TELEFAX: (202) 887-0763

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 517 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: peptide

PHOTHEICAL: NO

TI-SENSE: NO

IGINAL SOURCE:

INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13

73-089-30

Match 85.1%; Score 2582.5; DB 4; Length 517;

Local Similarity 91.4%; Pred. No. 3.le-266;

es 487; Conservative 7; Mismatches 8; Indels 31; Gaps 2;

18 SSERDCNNGEPPRKIIPEKNSURQYNSCARLCLNQETVCLASTAMKTCNCVAKTKLAN 77

13 ASERDCNDEPPRKIIITEKNTLRQ-----TKLAN 42

78 GTSSMIVPKQKLSASYEKELCVKYFEQWSESDQVEFVEHLISQMCHYQGHINSYLK 137

43 GTSSMIVPKQKLSANVEKELCVKYFEQWSESDQVEFVEHLISRMCHYQGHINTYLK 102

138 PMLORDFITALPARGLDHIAENILSYLDKSLCAELVCKEWYRVTSDGMLWKLIERMV 197

103 PMLORDFITALPARGLDHIAENILSYLDKSLCAELVCKEWYRVTSDGMLWKLIERMV 162

198 RTDSLWGLAERRGWGYLFKNKPPDGNAPPNSFYRALYPKIIDIETTESNWRGHRSL 257

163 RTDSLWGLAERRGWGYLFKNKPPDGNAPPNSFYRALYPKIIDIETTESNWRGHRSL 222

258 QRHCRSETSGKGYCLQYDDQKIVSGLRDNTIKWDKNTLECKRILTGHTGSLVCLQYDE 317

223 QRHCRSETSGKGYCLQYDDQKIVSGLRDNTIKWDKNTLECKRILTGHTGSLVCLQYDE 282

318 RVLIITGSSDSTVRVWDVNTGEMTLIIHCEAVLHLPNNGMVTCCKDRSIAVWDMASP 377

283 RVLIITGSSDSTVRVWDVNTGEMTLIIHCEAVLHLPNNGMVTCCKDRSIAVWDMASP 341

378 TDITLRRVLVGHRAAVNVVDFDKYIVSASGDRITIKVWNTSTCEFVRLTLNGHKGRIACIQ 437

342 TDITLRRVLVGHRAAVNVVDFDKYIVSASGDRITIKVWNTSTCEFVRLTLNGHKGRIACIQ 401

438 YRDLRVVSGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFNDKRIIVSGAYDGKIKVWDL 497

402 YRDLRVVSGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFNDKRIIVSGAYDGKIKVWDL 461

498 VAALDPRAPAGTICLRLTVBHSGRVRLQDFEFOIVSSSHDDTILWDFLNDP 550

462 VAALDPRAPAGTICLRLTVBHSGRVRLQDFEFOIVSSSHDDTILWDFLNDP 514

4

67-072A-30

nce 30, Application US/08487072A

t No. 6423684

RAL INFORMATION:

PLICANT: Mochly-Rosen, Daria

PLICANT: Ron, Dorit

TLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: Thereof

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morrison & Foerster

STREET: 2000 Pennsylvania Avenue, NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006-1812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/473,072A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: MURASHIGE, KATE H.

REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 2550-0025.20

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500

TELEFAX: (202) 887-0763

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 517 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: peptide

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13

US-08-487-072A-30

Query Match 85.1%; Score 2582.5; DB 4; Length 517;

Best Local Similarity 91.4%; Pred. No. 3.le-266;

Matches 487; Conservative 7; Mismatches 8; Indels 31; Gaps 2;

Qy 18 SSERDCNNGEPPRKIIPEKNSURQYNSCARLCLNQETVCLASTAMKTCNCVAKTKLAN 77

Db 13 ASERDCNDEPPRKIIITEKNTLRQ-----TKLAN 42

Qy 78 GTSSMIVPKQKLSASYEKELCVKYFEQWSESDQVEFVEHLISQMCHYQGHINSYLK 137

Db 43 GTSSMIVPKQKLSANVEKELCVKYFEQWSESDQVEFVEHLISRMCHYQGHINTYLK 102

Qy 138 PMLORDFITALPARGLDHIAENILSYLDKSLCAELVCKEWYRVTSDGMLWKLIERMV 197

Db 103 PMLORDFITALPARGLDHIAENILSYLDKSLCAELVCKEWYRVTSDGMLWKLIERMV 162

Qy 198 RTDSLWGLAERRGWGYLFKNKPPDGNAPPNSFYRALYPKIIDIETTESNWRGHRSL 257

Db 163 RTDSLWGLAERRGWGYLFKNKPPDGNAPPNSFYRALYPKIIDIETTESNWRGHRSL 222

Qy 258 QRHCRSETSGKGYCLQYDDQKIVSGLRDNTIKWDKNTLECKRILTGHTGSLVCLQYDE 317

Db 223 QRHCRSETSGKGYCLQYDDQKIVSGLRDNTIKWDKNTLECKRILTGHTGSLVCLQYDE 282

Qy 318 RVLIITGSSDSTVRVWDVNTGEMTLIIHCEAVLHLPNNGMVTCCKDRSIAVWDMASP 377

Db 283 RVLIITGSSDSTVRVWDVNTGEMTLIIHCEAVLHLPNNGMVTCCKDRSIAVWDMASP 341

Qy 378 TDITLRRVLVGHRAAVNVVDFDKYIVSASGDRITIKVWNTSTCEFVRLTLNGHKGRIACIQ 437

Db 342 TDITLRRVLVGHRAAVNVVDFDKYIVSASGDRITIKVWNTSTCEFVRLTLNGHKGRIACIQ 401

Qy 438 YRDLRVVSGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFNDKRIIVSGAYDGKIKVWDL 497

Db 402 YRDLRVVSGSSDNTIRLWDIECGACLRVLEGHEELVRCIRFNDKRIIVSGAYDGKIKVWDL 461

Match 13.2%; Score 399; DB 3; Length 779;
 Local Similarity 24.3%; Pred. No. 3.8e-33;
 es 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

56 TVCLASTAMKTCVAKTKLANGTSSMIVPKQKLSASYEK-----ELCVKVFQW 108
 194 TTPLAKTTKINN-----NNNIADLIESKOSIISPEYLSDEIFSAINNPLHAYFK-- 244

109 SESQVFEVHLISQCHYQHINSYLPKMLQORDITLALPARGLDHIAENILSYLDKAS 168
 245 -----NLFLVANNDRSLSDLGLTKDKNLKDLITSLPF-----EISLKFNYLQFED 294

169 LCAAEVCKEWEYR-VTSDGMLWKKLI--ERMVTDLSLWGLAERGGWGYLFKNKPPDGN 225
 295 IINSLGVSQWNKIIIRKSTLSMKLLISENFV-----SPKGF 331

226 APPNSFYRALYPKIIQD-----IET--IESNWRGCRHSLQRIHCRSETSKGVYCLQYD 276
 332 NSLNKLSQKYPKLSQDRLRLSFLNIFILKNWYNPKFVPORTTLRGHTSVITCLOFE 391

277 DQKIVGLRNTIKIWDKNTLECKRILTGHTSVLCLOYDE-RVITGSSDSTVRWVDN 335
 392 DNYVITGADDKMIRVYDSINKKFLQLSGHDGGWALKYAHGGILVSGSDTRVWVDIK 451

336 TGEMLNTLIHCEAVLHLRFNNGMMVTCCKDRSIAVWDMASPTDITLRVLVGHRAAVNV 395
 452 KG-----CCT-----HVFEGHSTVRC 468

396 VDFDD---KYIVSASGDRITIKVWNT-----STCE-----FVRLTNG 428
 469 LDIVEYKNIKYIVTGSRDNTLHWKLPKSSVPDHGEEHDYPLVFTHTPEENPFVGVLRG 528

429 HKRGIAQLQVDRDLVVGSSDNTIRLWDIECGACLRVLEGHVELVRCIRFDN--KRIVSG 486
 529 HMASVTVSGHGNIVVSGSYDNTLIVMDVAQMKCLYLSGHTDRIYSTIYDHERKRCISA 588

487 AYDGKIKVMDL-----VAALDPRAPAGTL--CLRTLVEHSGRVFRLQDFEQIVSS 535
 589 SMDTTIRIWDLENINWNGECSYATNSGPCAKILGMWYTLQGHIALVGLLRUSDKFLVSA 648

536 SHDDTILIMDFLNDPAAQAEPPRSPRTYTY 566
 649 AADGSIRGWD-AND-----YSRKFSY 668

9

73-089-32
 nce 32, Application US/08473089
 t No. 6342368
 RAL INFORMATION:
 PLICANT: Mochly-Rosen, Daria
 PLICANT: Ron. Borit
 ILE OF INVENTION: WD-40 - Derived Peptides and Uses
 ILE OF INVENTION: Thereof
 MBER OF SEQUENCES: 265
 RRESPONDENCE ADDRESS:
 ADDRESSEE: Morrison & Foerster
 STREET: 2000 Pennsylvania Avenue, NW
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812
 4PUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 RENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/473,089
 ILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 CORNEY/AGENT INFORMATION:

NAME: MURASHIGE, KATE H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 2550-0025.22
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 INFORMATION FOR SEQ ID NO: 32:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 779 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: CDC4 / CDC20 protein, Fig. 15
 US-08-473-089-32

Query Match 13.2%; Score 399; DB 4; Length 779;
 Best Local Similarity 24.3%; Pred. No. 3.8e-33;
 Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

Qy 56 TVCLASTAMKTCVAKTKLANGTSSMIVPKQKLSASYEK-----ELCVKVFQW 108
 Db 194 TTPLAKTTKINN-----NNNIADLIESKOSIISPEYLSDEIFSAINNPLHAYFK-- 244

Qy 109 SESQVFEVHLISQCHYQHINSYLPKMLQORDITLALPARGLDHIAENILSYLDKAS 168
 Db 245 -----NLFLVANNDRSLSDLGLTKDKNLKDLITSLPF-----EISLKFNYLQFED 294

Qy 169 LCAAEVCKEWEYR-VTSDGMLWKKLI--ERMVTDLSLWGLAERGGWGYLFKNKPPDGN 225
 Db 295 IINSLGVSQWNKIIIRKSTLSMKLLISENFV-----SPKGF 331

Qy 226 APPNSFYRALYPKIIQD-----IET--IESNWRGCRHSLQRIHCRSETSKGVYCLQYD 276
 Db 332 NSLNKLSQKYPKLSQDRLRLSFLNIFILKNWYNPKFVPORTTLRGHTSVITCLOFE 391

Qy 277 DQKIVGLRNTIKIWDKNTLECKRILTGHTSVLCLOYDE-RVITGSSDSTVRWVDN 335
 Db 392 DNYVITGADDKMIRVYDSINKKFLQLSGHDGGWALKYAHGGILVSGSDTRVWVDIK 451

Qy 336 TGEMLNTLIHCEAVLHLRFNNGMMVTCCKDRSIAVWDMASPTDITLRVLVGHRAAVNV 395
 Db 452 KG-----CCT-----HVFEGHSTVRC 468

Qy 396 VDFDD---KYIVSASGDRITIKVWNT-----STCE-----FVRLTNG 428
 Db 469 LDIVEYKNIKYIVTGSRDNTLHWKLPKSSVPDHGEEHDYPLVFTHTPEENPFVGVLRG 528

Qy 429 HKRGIAQLQVDRDLVVGSSDNTIRLWDIECGACLRVLEGHVELVRCIRFDN--KRIVSG 486
 Db 529 HMASVTVSGHGNIVVSGSYDNTLIVMDVAQMKCLYLSGHTDRIYSTIYDHERKRCISA 588

Qy 487 AYDGKIKVMDL-----VAALDPRAPAGTL--CLRTLVEHSGRVFRLQDFEQIVSS 535
 Db 589 SMDTTIRIWDLENINWNGECSYATNSGPCAKILGMWYTLQGHIALVGLLRUSDKFLVSA 648

Qy 536 SHDDTILIMDFLNDPAAQAEPPRSPRTYTY 566
 Db 649 AADGSIRGWD-AND-----YSRKFSY 668

RESULT 10
 US-08-487-072A-32
 ; Sequence 32, Application US/08487072A
 ; Patent No. 6423684
 ; GENERAL INFORMATION:
 ; APPLICANT: Mochly-Rosen, Daria
 ; APPLICANT: Ron. Borit
 ; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 ; TITLE OF INVENTION: Thereof
 ; NUMBER OF SEQUENCES: 265

RESPONSE ADDRESS:

ADDRESSEE: Morrison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA

ZIP: 20006-1812

MPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,072A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514

TORNEY/AGENT INFORMATION:

NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2550-0025.20
LECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
RATION FOR SEQ ID NO: 32:
QUENCE CHARACTERISTICS:
LENGTH: 779 amino acids
TYPE: amino acid
TOPOLOGY: unknown
LECULE TYPE: protein
POTHEICAL: NO
TI-SENSE: NO
IGNAL SOURCE:
INDIVIDUAL ISOLATE: CDC4 / CDC20 protein, Fig. 15
87-072A-32

Match 13.2%; Score 399; DB 4; Length 779;

Local Similarity 24.3%; Pred. No. 3.8e-33;
es 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

56 TVCLASTAMKTEVCVAKTKLANGTSSMIVPKQKLSASYEKEK-----ELCVKYFEOW 108

194 TTPLAKTKTINN-----NNNIADIESKDSIIISPEYLSDEIFSAINNPLHAYFK-- 244

109 SESDQVEFVHLISQMHCHYGHINSYLKPEMLQORDITAPARGLDHIAENILSYLDAS 168

245 -----NLLPRLVANMDRSELSDLGLTIKDNKREDLITSLPF-----EISLKIFNYLOFED 294

169 LCAAEVLCKEWR--VTS DGMWKLI--ERMVTDLSLRGLAERRGWGYLFKNKPPDGN 225

295 IINSLGVSONWNIIRKSTSLWKLLISENFV-----SPKGF 331

226 APPNSFYRALYPKIID-----IET--IESNRCGRHSLORIHCRSETSKGVYCLQYD 276

332 NSLNLKLSQKYPKLSQDRLRLSFLNIFILKWNYPKFPVQRTTLRGHMTSVITCLOFE 391

277 DQKIVSLRNTIKINDKNTLECKRILTGTGVLCLQYDE-RVIITGSSDSTVRVMDVN 335

392 DNYVITGADDKMIRVYDSINKKELLQSGHGGVWALKYAHGGILVSGSTDRVYWDIK 451

336 TGEMLNTLIHCEAVLHLRFNNGMWVTCSDRSIAVWDNASPTDITLRLVVGHRVAVNV 395

452 KG-----CCT-----HVFEGHNSTVRC 468

396 VFDD-----KYIVSAGDRTIKVWNT-----STCE-----FVRLTNG 428

469 LDIVEYKNIYVTGSRDNTLHVWKLPKESSVPDHGEEDYPLVPHTPPENPFVGVLRG 528

429 HKRGIAQLQVDRLVVSGSDNTIRLWDIECGACLAVLEGHEELVRCIFDN--KRVISG 486

529 HMASVRTVSGHGNIVVSGSYDNTLIVMDVAQMKCLVILSGHTDRIVSTIYDHERKRCISA 588

487 AYDGKIKVWDL-----VAALDPRAPAGTL--CLRTLVEHSGRVFLQDFEFOIVSS 535

Db 589 SMDTIRIWDLNINWNGECSYATNSASPCAKILGAMYTLOGHTALVGLRLSLDKFLVSA 648

Qy 536 SHDPTILWDLNDPAQAEPSPSRPTY 566

Db 649 AADGSGIRGWD-AND-----YSRKFSY 668

RESULT 11

US-09-177-165A-29

; Sequence 29, Application US/09177165A

; Patent No. 6426205

; GENERAL INFORMATION:

; APPLICANT: Tyers, Mike

; APPLICANT: Willems, Andrew

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBIQUITIN

; TITLE OF INVENTION: DEPENDENT PROTEOLYSIS

; FILE REFERENCE: 11757.10USUI

; CURRENT APPLICATION NUMBER: US/09/177,165A

; CURRENT FILING DATE: 1998-10-22

; PRIOR APPLICATION NUMBER: 60/092,443

; FILING DATE: 1998-07-10

; PRIOR APPLICATION NUMBER: 60/063,254

; PRIOR FILING DATE: 1997-10-24

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 29

; LENGTH: 779

; TYPE: PRT

; ORGANISM: Saccharomyces cerevisiae

US-09-177-165A-29

Query Match 13.2%; Score 399; DB 4; Length 779;

Best Local Similarity 24.3%; Pred. No. 3.8e-33;

Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

Qy 56 TVCLASTAMKTEVCVAKTKLANGTSSMIVPKQKLSASYEKEK-----ELCVKYFEOW 108

Db 194 TTPLAKTKTINN-----NNNIADIESKDSIIISPEYLSDEIFSAINNPLHAYFK-- 244

Qy 109 SESDQVEFVHLISQMHCHYGHINSYLKPEMLQORDITAPARGLDHIAENILSYLDAS 168

Db 245 -----NLLPRLVANMDRSELSDLGLTIKDNKREDLITSLPF-----EISLKIFNYLOFED 294

Qy 169 LCAAEVLCKEWR--VTS DGMWKLI--ERMVTDLSLRGLAERRGWGYLFKNKPPDGN 225

Db 295 IINSLGVSONWNIIRKSTSLWKLLISENFV-----SPKGF 331

Qy 226 APPNSFYRALYPKIID-----IET--IESNRCGRHSLORIHCRSETSKGVYCLQYD 276

Db 332 NSLNLKLSQKYPKLSQDRLRLSFLNIFILKWNYPKFPVQRTTLRGHMTSVITCLOFE 391

Qy 277 DQKIVSLRNTIKINDKNTLECKRILTGTGVLCLQYDE-RVIITGSSDSTVRVMDVN 335

Db 392 DNYVITGADDKMIRVYDSINKKELLQSGHGGVWALKYAHGGILVSGSTDRVYWDIK 451

Qy 336 TGEMLNTLIHCEAVLHLRFNNGMWVTCSDRSIAVWDNASPTDITLRLVVGHRVAVNV 395

Db 452 KG-----CCT-----HVFEGHNSTVRC 468

Qy 396 VFDD-----KYIVSAGDRTIKVWNT-----STCE-----FVRLTNG 428

Db 469 LDIVEYKNIYVTGSRDNTLHVWKLPKESSVPDHGEEDYPLVPHTPPENPFVGVLRG 528

Qy 429 HKRGIAQLQVDRLVVSGSDNTIRLWDIECGACLAVLEGHEELVRCIFDN--KRVISG 486

Db 529 HMASVRTVSGHGNIVVSGSYDNTLIVMDVAQMKCLVILSGHTDRIVSTIYDHERKRCISA 588

Qy 487 AYDGKIKVWDL-----VAALDPRAPAGTL--CLRTLVEHSGRVFLQDFEFOIVSS 535

Db 589 SMDTIRIWDLNINWNGECSYATNSASPCAKILGAMYTLOGHTALVGLRLSLDKFLVSA 648

Qy 536 SHDPTILWDLNDPAQAEPSPSRPTY 566

Gaps 8;


```

280 IVSGLRNTTIKIWDKNTLCKKILTLGHTSGVLCLOYDE--RVIITGSSDSTVRVWDVNTG 337
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
122 MVSASEDATIKWMDYETGDFEXTLKHTSDVSODIFSDHSGLLKASCADMTIKLWDFQGF 181
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
338 EMLNTLIHHCCEAVLHURF--NNGMMVTCSKORSIAVMDMASPTDITLRRVLGVGHRAAVNV 395
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
182 ECIRTMHGHDHNVSSVAIWPNGDHIVSASRUKTIKWHEVQTGYCV---KIFTGHERWEMV 238
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
396 V--DFDDKIYIVSAGSDRTTIKWNTSTCEPVTTLNGHKRGIACTOYDR-----441
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
239 VRPNQDGLIASNSDQTVRVVWVATKCEABLRHEHWEVCEISWAPESSYSSISEATGS 298
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
442 -----LVVSGSDNTIRLWDIECGACLRVLGHEELVRCIRDN---KRIVSGAYD 489
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
299 ETKSGXPGFILSGSRDKTIKNWDVSTGCMUMLTVGHNDNVRGVLPHSGCKFILSCAD 358
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
490 GKIKVLDVAALDPRAPAGTLCRLTVLHBSGRVPRPQLQDEF--QIVSSSHDDTILWID 545
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
359 KTLRVWDY-----KNKRCWKTLNAHEHFVSLDPKHTAPVVTGTVSDVTKVWE 407
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

14 961-716-3
Application 3, Application US/08961716
at No. 580272
SERIAL INFORMATION:
APPLICANT: ADACHI, HIDEKI
APPLICANT: TSUJIMOTO, MASAFUMI
APPLICANT: INOUE, KEIZO
APPLICANT: ARAI, HIROYUKI
TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME
TITLE OF INVENTION: AND GENE THEREOF
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER &
ADDRESSEE: NEUSTADT, P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,716

FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/283,917
 FILING DATE: 03-AUG-1994
 APPLICATION NUMBER: JP 205943/1993
 FILING DATE: 03-AUG-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: ObiLen, NO. 5880272man F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 2292-030-0
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 413-3000
 TELEFAX: (703) 413-2220
 TELEX: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 409 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 LECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Bos taurus
 61-716-3

```

Query Match      11.7%; Score 354; DB 2; Length 409;
Best Local Similarity 29.2%; Pred. No. 8e-29;
Matches      87; Conservative 57; Mismatches 110; Indels 44; Gaps 8;

QY      280  IVSGURDNTIKIKWDKNTLCKRILTGHTGSLVLCLOYDE--RVIIITGSSDSTVRVMDVNTG 337
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      122  MVASGEDATIKVWDYETGPFERTLKGHTDSVQDISFDHSGKLLASCSADMTIKLWDFQGF 181

QY      338  EMLNTLIHCEAVLHLP--NNGMMVTSKORSIAVWDMASFTDITLRVLVGHRAAVNV 395
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      182  ECIRTMGHGDHNVSSVAINPNGDHIVSASRDKTIKNWEVOTGYCV--KTFTGHEWVRM 238

QY      396  V--DPEDKVIIVSASGDRTIKVWNITSTCFVRTLNGHKRGIAQLQYRDR----- 441
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      239  VRPNQDGTLIASCNSDQTVRVVWVATKECKAELREHEHVVEICISWAPESSYSSISEATGS 298

QY      442  -----LVVSGSSDNTILWDIECGALRVLEGHEELVRCIRFDN--KRVISGAYD 489
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      299  EYTKSGKPGPFLLVSGSRDKTIKNWDMGTGCMCLTVGHNNVRGVLTFHSGGKPFILSCADD 358

QY      490  GKIKWDLVAALDPRAPAGTLCRLTIVHSGRVFLQFDEF--QIVSSSHDDTILIWD 545
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      359  KTLRVWDY-----KNKRCKMNLNAHEHFTSLDFHKTPAPYVVTGSDVOTKWE 407
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

RESULT 15
 US-08-283-917-9
 Sequence 9, Application US/08283917
 Patent No. 5849557
 GENERAL INFORMATION:
 APPLICANT: ADACHI, HIDEKI
 APPLICANT: TSUJIMOTO, MASAFUMI
 APPLICANT: INOUE, KEIZO
 APPLICANT: ARAI, HIROYUKI
 TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME
 TITLE OF INVENTION: AND GENE THEREOF
 NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &
 ADDRESSEE: NEUSTADT,P.C.
 STREET: 1755 S. Jefferson Davis Highway, Suite 400
 City: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/283,917
 FILING DATE: 03-AUG-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 209943/1993
 FILING DATE: 03-AUG-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Oblon, No. 5849557man F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 2292-030-0
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 413-3000
 TELEFAX: (703) 413-2220
 TELEX: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 410 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-283-917-9

Y Match 11.7%; Score 354; DB 2; Length 410;
Local Similarity 29.2%; Pred. No. 8.1e-29;
hes 87; Conservative 57; Mismatches 110; Indels 44; Gaps 8;
280 IVSGLRDNTIKIWKNTLECKRILTGHTGTVLCLOYDE--RVITGSSDSTVRVWDVNTG 337
123 MVSASEDAIKVWDYETGDFERTLKHTDSVQDISFDHSGKLLASCADMTIKLWDFQGF 162
338 EMLNTLIHCEAVLHREF--NNGMMVTCKDRSIADVMDASPTDITLRRVLVGHRAAVNV 395
183 ECIRTMHGHDNVSSVAIMPNGDHIYSASRDXTIKMWEVQTGYCV---KTFTGHREWRM 239
396 V--DFDDKYIIVSASGDRTIKVNWTSTCEVFRTLNHKGKGIACLOYRDR----- 441
240 VRPNQDGTLIASCNSNDQTVRYVWVATKECAELREHEHVVECIWAPESYSISEATGS 239
442 -----LVVSGSDNTIRLWDIECGACLRVLEGHEELVRCIRFDN--KRIVSGAYD 489
300 ETKKSGKPGPFLLSGSRDXTIKMWDVSTGCMCLMTLVGHDNWRGVLFHSGGKFILSCADD 359
490 GKIKVWDLVAALDPRAPAGTLCRLTVEHSGRVFRLOPDEF--QIVSSSHDDTILIWD 545
360 KTLRWVDY-----KNKRCMKTLANAHEHFTSLDPFKTAPYVVTGSDVDTVKVME 408

completed: October 22, 2003, 10:26:37
me : 32 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

ein - protein search, using sw model

October 22, 2003, 10:14:07 ; Search time 104 Seconds
(without alignments)
1411.845 Million cell updates/sec

US-09-601-168b-2

score: 3034

e: 1 MDPAEAVLQEKALKEFMSSE.....PAQAEPSPSRPTTYISR 569

table: BLOSUM62

Gapop 10.0 , Gapext 0.5

d: 830525 seqs, 258052604 residues

umber of hits satisfying chosen parameters: 830525

DB seq length: 0

DB seq length: 2000000000

rocessing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

e: SPTREMBL_23:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mmc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archheap:*

red. No. is the number of results predicted by chance to have a
core greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Score	Query Match	Length	DB	ID	Description
2397	98.8	569	11	Q9QUI5	Q9qui5 mus musculus
2390	98.5	569	11	Q9Z159	Q9z159 mus musculus
2378	98.2	569	11	Q9R1G7	Q9r1g7 mus musculus
2395	78.9	529	11	Q8BY90	Q8by90 mus musculus
2380	78.4	563	11	Q8K022	Q8k022 mus musculus
2360	77.8	555	11	Q8CHD5	Q8chd5 mus musculus
2386	75.3	563	11	Q9Z3H0	Q9z3h0 mus musculus
2089.5	68.9	510	5	O44382	O44382 drosophila
2089.5	68.9	510	5	Q9VDE3	Q9vde3 drosophila
1498	49.4	430	5	Q9BJ54	Q9bj54 heterodera
1230	40.5	252	11	Q9Z2C7	Q9z2c7 mus musculus
661	21.8	1326	5	Q9VZF4	Q9vzf4 drosophila
640	21.1	553	4	Q9NUX6	Q9nux6 homo sapien
640	21.1	561	4	Q96R12	Q96r12 homo sapien
640	21.1	589	4	Q96LE0	Q96le0 homo sapien
640	21.1	627	4	Q96A16	Q96a16 homo sapien

Q969h0 homo sapien
Q8vhp4 mus musculu
Q8vbw4 mus musculu
Q8uun3 xenopus lae
Q95zt0 caenothabdi
Q44083 caenothabdi
Q8ccs5 mus musculu
Q8xlp2 podospora a
Q8xlp4 podospora a
Q8xlp3 podospora a
Q8xlp5 podospora a
Q8z0r1 anabaena sp
Q8uun8 xenopus lae
Q8d4t2 mus musculu
Q8n136 homo sapien
Q8n776 homo sapien
Q8y109 anabaena sp
Q8z020 anabaena sp
Q96611 dictyosteli
Q8vz19 anabaena sp
Q8ysc0 anabaena sp
Q90z14 xenopus lae
Q9ptr5 gallus gall
Q8hxx0 macaca fasc
Q9gl51 sus scrofa
Q8avn1 xenopus lae
Q8xb12 chlorobium
Q96698 drosophila
Q810f4 dictyosteli

ALIGNMENTS

RESULT 1

Q9QUI5 PRELIMINARY; PRT; 569 AA.
AC Q9QUI5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Ubiquitin ligase FWD1 (Beta-transducin repeat containing protein)
DE (F-box-WD40 repeat protein 1).
GN BTBC OR FBXW1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9919275; PubMed=10097128;
RA Hatakeyama S., Kitagawa M., Nakayama K., Shirane M., Matsumoto M.,
RA Hattori K., Higashi H., Nakano H., Okumura K., Onoe K., Good R.A.,
RA Nakayama K.-I.;
RT "Ubiquitin-dependent degradation of IkappaBalpha is mediated by a
RT ubiquitin ligase Skp1/Cul 1/F-box protein FWD1.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:3859-3863(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99075339; PubMed=9859996;
RA Yaron A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,
RA Andersen J.S., Mann M., Mercuro F., Ben-Neriah Y.;
RT "Identification of the receptor component of the IkappaBalpha-
RT ubiquitin ligase.";
RL Nature 396:590-594(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX Strausberg R.;
RA Submitted (FEF-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=21601157; PubMed=11735228;

uyama S., Hatakeyama S., Nakayama K., Ishida N., Kawakami K.,
Kawaya K.-I.;
Characterization of a Mouse Gene (Fbxw6) That Encodes a Homologue of
the Drosophila Protein F-box-10.
Genomics 78:214-222(2001).
AF081887; AAD08701.1; -
AF099932; AAD08701.1; -
BC003989; AAH03989.1; -
AF391190; AAL40329.1; -
AF391178; AAL40329.1; JOINED.
AF391179; AAL40329.1; JOINED.
AF391180; AAL40329.1; JOINED.
AF391181; AAL40329.1; JOINED.
AF391182; AAL40329.1; JOINED.
AF391183; AAL40329.1; JOINED.
AF391184; AAL40329.1; JOINED.
AF391185; AAL40329.1; JOINED.
AF391186; AAL40329.1; JOINED.
AF391187; AAL40329.1; JOINED.
AF391188; AAL40329.1; JOINED.
AF391189; AAL40329.1; JOINED.
AF391190; AAL40329.1; JOINED.
D1; MGI:1338871; Btcr.
IPR001810; P-box.
IPR001680; WD40.
PF00645; F-box; 1.
PF00400; WD40; 7.
Pfam; PRO0320; GPROTEINERPT.
ProDom; PD000018; WD40; 4.
SMART; SM00256; FBOX; 1.
SMART; SM00320; FBOX; 7.
SITE; PS00181; FBOX; 1.
SITE; PS00678; WD_REPEATS_1; 6.
SITE; PS00082; WD_REPEATS_2; 7.
SITE; PS50294; WD_REPEATS_REGION; 1.
Repeat; WD repeat.
SEQUENCE 569 AA; 65105 MW; BC7D6544815B2296 CRC64;

Match 98.8%; Score 2997; DB 11; Length 569;
Local Similarity 98.8%; Pred.No. 9.6e-249;
Gaps 56; Conservative 3; Mismatches 5; Indels 0; Gaps 0

1 MDPAEAVLQEKALFKWNSSEREDCNNGEPKKIIPEKNSLRQTYSNCSARCLNQETVCLIA 60
1 MDPAEAVLQEKALFKWNSSEREDCNNGEPKKIIPEKNSLRQTYSNCSARCLNQETVCLUT 60
61 STAMKTENCYAKTKLANGTSSMIVPKORLSASYEKEKELCVKFEQWSSSDQVEFVEHL 120
61 STAMKTENCYAKTKLANGTSSMIVPKORLSASYEKEKELCVKFEQWSSSDQVEFVEHL 120
121 ISOMCHYQHGHINSYLKPMLQRDFITAPARGLDHIAENILSYLDAKSLSAAELVCWKWY 180
121 ISOMCHYQHGHINSYLKPMLQRDFITAPARGLDHIAENILSYLDAKSLSAAELVCWKWY 180
181 RVTSDBGWLWKLIERWVRTDSLWRGLAERGQGVLFNKNKPDGNAPNSFYALVPKLI 240
181 RVTSDBGWLWKLIERWVRTDSLWRGLAERGQGVLFNKNKPDGNAPNSFYALVPKLI 240
241 QDIETIESNRWCRGHSRLQIRHCSETSKGVYCLQYDDQKIIVSGLRDNTIKWDKSTLECK 300
241 QDIETIESNRWCRGHSRLQIRHCSETSKGVYCLQYDDQKIIVSGLRDNTIKWDKSTLECK 300
301 RILTGTGTVLCLOYDERVLIITGSSDSVTVMVDVNTGBMLNTLIHHCEAVLHFRFNGMW 360
301 RILTGTGTVLCLOYDERVLIITGSSDSVTVMVDVNTGBMLNTLIHHCEAVLHFRFNGMW 360
361 VTCSKDRSIAVMDMASPTDITLRRVLVGHRAAVVVDFDDKYIVSASGDRTIKVWNTSTC 420
361 VTCSKDRSIAVMDMASPTDITLRRVLVGHRAAVVVDFDDKYIVSASGDRTIKVWNTSTC 420
421 EFVETLNCHKRGACIQYDRDLVVGSSDNTRLWDIECGACLVRLEGHELVRCIRFON 480
421 EFVETLNCHKRGACIQYDRDLVVGSSDNTRLWDIECGACLVRLEGHELVRCIRFON 480

Qy	481	KNIVSGAVDGKIKWDLVAALDPRAPAGTLCRLTLVHSGRVRPRLQDFBQFQIVSSSHDDT	540			
Db	481	KNIVSGAVDGKIKWDLVAALDPRAPAGTLCRLTLVHSGRVRPRLQDFBQFQIVSSSHDDT	540			
Qy	541	ILIWFLNDPAAQAEPSPRSRTYTIYSR	569			
Db	541	ILIWFLNDPAAHAEPSPRSRTYTIYSR	569			
RESULT 2						
ID	Q9Z1159	PRELIMINARY; PRT; 569 AA.				
AC	Q9Z159;					
DT	01-MAY-1999	(TrEMBLrel. 10, Created)				
DT	01-MAY-1999	(TrEMBLrel. 10, Last sequence update)				
DT	01-OCT-2002	(TrEMBLrel. 22, Last annotation update)				
DE	Beta-transducin repeat containing protein.					
GN	BTRC.					
OS	Mus musculus (Mouse).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
OX	NCBI_TaxID=10090;					
FN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=99145465; PubMed=9990853;					
RA	Spencer E., Jiang J., Chen Z.J.;					
RT	"Signal-induced ubiquitination of IkappaBalpha by the F-box protein					
RT	Slimo/beta-Trcp.";					
RL	Genes Dev. 13:284-294(1999).					
DR	EMBL; AF112979; AAD04181.1; -.					
DR	MGD; MG1:1338871; BTRC.					
DR	InterPro; IPR001810; F-box.					
DR	InterPro; IPR001680; WD40.					
DR	Pfam; PF00646; F-box; 1.					
DR	Pfam; PF00400; WD40; 7.					
DR	PRINTS; PR00320; GPROTEINBRPT.					
DR	ProDom; PD000018; WD40; 4.					
DR	SMART; SM00256; FBOX; 1.					
DR	SMART; SM00320; WD40; 7.					
DR	PROSITE; PS0181; FBOX; 1.					
DR	PROSITE; PS00678; WD REPEATS 1; 6.					
DR	PROSITE; PS00882; WD REPEATS 2; 7.					
DR	PROSITE; PS0294; WD REPEATS_REGION; 1.					
KW	Repeat; WD repeat.					
SQ	SEQUENCE 569 AA; 65047 MW; BC7C7A44815BED96 CRC64;					
Query Match 98.5%; Score 2990; DB 11; Length 569;						
Best Local Similarity 98.4%; Pred. No. 3 8e-248;						
Matches 560; Conservative 3; Mismatches 6; Indels 0; Gaps 0;						
Qy	1	MDPAVAILOEKALXFMNSSREDCCNNGEPPKRIPEKNSLRQTYNSCARLCLNQETVCLA	60			
Db	1	MDPAVAILOEKALXFMNSSREDCCNNGEPPKRIPEKNSLRQTYNSCARLCLNQETVCLT	60			
Qy	61	STAMKTENCVAKTLANGTSSMTVPKOKLSAYEKEKLCVKYFEQWSESDQVEFVEHL	120			
Db	61	STAMKTENCVAKALANGTSSMTVPKOKLSAYEKEKLCVKYFEQWSESDQVEFVEHL	120			
Qy	121	ISQMCHYQRGHINSYLPMLQRFITPALPARGLDHIAENILSYLDAKSLCAELVCKEWY	180			
Db	121	ISQMCHYQRGHINSYLPMLQRFITPALPARGLDHIAENILSYLDAKSLCAELVCKEWY	180			
Qy	181	RVTSQGLMKKLIERWVRTSLWRGLAERGWQYILFKNKPQGNAPPNSFYRALYPKII	240			
Db	181	RVTSQGLMKKLIERWVRTSLWRGLAERGWQYILFKNKPQGNAPPNSFYRALYPKII	240			
Qy	241	QDIETIENWRGHSRQRHCSRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK	300			
Db	241	QDIETIENWRGHSRQRHCSRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK	300			
Qy	301	RILTGHTGSVLCLOQYDERVITIGSSDSTVRVWDVNTGEMLNTLIHCEAVLHLRFNNGMM	360			
Db	301	RILTGHTGSVLCLOQYGERVITIGSSDSTVRVWDVNTGEMLNTLIHCEAVLHLRFNNGMM	360			

```

361 VTSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDDKXIVSASGDRITIKVNTSTC 420
|||||
361 VTSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDDKXIVSASGDRITIKVNTSTC 420
|||||
421 EFVRLTNGHKGRIACIQYRDRLVVGSSDNTIRLWDIECGACLRVLEGGHEELVRCIRFDN 480
|||||
421 EFVRLTNGHKGRIACIQYRDRLVVGSSDNTIRLWDIECGACLRVLEGGHEELVRCIRFDN 480
|||||
481 KRIVSGAYDGKIKWDLVAALDPRAPAGTLCRLTLVEHSGRVFRLQDFEQIVSSSHDDT 540
|||||
481 KRIVSGAYDGKIKWDLVAALDPRAPAGTLCRLTLVEHSGRVFRLQDFEQIVSSSHDDT 540
|||||
541 ILIWDFLNDPAAOAEPPRSPRTYTIYSR 569
|||||
541 ILIWDFLNDPAAOAEPPRSPRTYTIYSR 569
|||||

```

3

R1G7 PRELIMINARY; PRT; 569 AA.

```

-MAY-2000 (TrEMBLrel. 13, Created)
-MAY-2000 (TrEMBLrel. 13, Last sequence update)
-OCT-2002 (TrEMBLrel. 22, Last annotation update)
ta-transducin repeat-containing protein.
"RC.

```

```

us musculus (Mouse).
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
"BI_TaxID=10090;
}

```

```

SEQUENCE FROM N.A.
nston J., Ellledge S.J., Harper J.W.;
bmitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
BL; AF10396; AA041025.1; -.
D; MG1:1338871; Btrc.
terPro; IPR001810; F-box.
am; PF00646; F-box; 1.
am; PF00400; WD40; 7.

```

INTS; PR00320; GPROTEINRPT.

CDOM; PD000018; WD40; 4.

ART; SM00256; FBOX; 1.

ART; SM00320; WD40; 7.

OSITE; PS50181; FBOX; 1.

OSITE; PS00678; WD_REPEATS_1; 6.

OSITE; PS00082; WD_REPEATS_2; 7.

OSITE; PS00294; WD_REPEATS_REGION; 1.

peat; WD repeat.

QUENCE 569 AA; 65209 MW; E6DDCAD28D551D9D CRC64;

Match 98.2%; Score 2978; DB 11; Length 569;

Local Similarity 97.9%; Pred. No. 4.1e-247;

es 557; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

1 MDPAEAVLQEKALFPMNSRERDCNNGEPKRIPEKNSLRQYNSCARLCLNQETVCLA 60

1 MDPAEAVLQEKALFPMNSRERDCNNGEPKRIPEKNSLRQYNSCARLCLNQETVCLT 60

61 STAMKTCNCVAKTKLANGTSSMIVPKORKLSASYEKELCVKYFEQWSSDDQVEFVEHL 120

61 STAMKTCNCVAKTKLANGTSSMIVPKORKLSASYEKELCVKYFEQWSSDDQVEFVEHL 120

121 ISQCHYQHGHSINLYKPLMLQORDPITALPARGLDHIAENILSYLDAKSLCAAELVCKEY 180

121 ISQCHYQHGHSINLYKPLMLQORDPITALPARGLDHIAENILSYLDAKSLCAAELVCKEY 180

181 RVTSDGMLWKKLIERMVRLTSLMRGLAERRGWGYLFKNKPPDGNAPPNSFYRALKPKII 240

181 RVTSDGMLWKKLIERMVRLTSLMRGLAERRGWGYLFKNKPPDGNAPPNSFYRALKPKII 240

241 QDIETIESNRCGRHSQRHCRSETSGKGYVCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300

```

Db 241 QDIETIESNRCGRHSQRHCRSETSGKGYVCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300
|||||
Qy 301 RILTGHGTVLCLOYDERVITGSSDSTVRWVDMVTGEMNTLIHHCCEAVLHFRFNNGMM 360
|||||
Db 301 RILTGHGTVLCLOYDERVITGSSDSTVRWVDMVTGEMNTLIHHCCEAVLHFRFNNGMM 360
|||||
Qy 361 VTSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDDKXIVSASGDRITIKVNTSTC 420
|||||
Db 361 VTSKDRSIAVWDMASPTDITLRRVLVGHRAAVNVDFDDKXIVSASGDRITIKVNTSTC 420
|||||
Qy 421 EFVRLTNGHKGRIACIQYRDRLVVGSSDNTIRLWDIECGACLRVLEGGHEELVRCIRFDN 480
|||||
Db 421 EFVRLTNGHKGRIACIQYRDRLVVGSSDNTIRLWDIECGACLRVLEGGHEELVRCIRFDN 480
|||||
Qy 481 KRIVSGAYDGKIKWDLVAALDPRAPAGTLCRLTLVEHSGRVFRLQDFEQIVSSSHDDT 540
|||||
Db 481 KRIVSGAYDGKIKWDLVAALDPRAPAGTLCRLTLVEHSGRVFRLQDFEQIVSSSHDDT 540
|||||
Qy 541 ILIWDFLNDPAAOAEPPRSPRTYTIYSR 569
|||||
Db 541 ILIWDFLNDPAAOAEPPRSPRTYTIYSR 569
|||||

```

RESULT 4

```

Q8BY90 PRELIMINARY; PRT; 529 AA.
AC Q8BY90;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE F-box/WD-repeat protein 1B.
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).
DR EMBL: AK041532; BAC30975.1; -.
SQ SEQUENCE 529 AA; 60812 MW; D2F382457FD90080 CRC64;

```

```

Query Match 78.9%; Score 2395; DB 11; Length 529;
Best Local Similarity 78.9%; Pred. No. 4.6e-197;
Matches 449; Conservative 43; Mismatches 37; Indels 40; Gaps 3;

```

```

Qy 1 MDPAEAVLQEKALFPMNSRERDCNNGEPKRIPEKNSLRQYNSCARLCLNQETVCLA 60
|||||
Db 1 MEP-DSVIEDKTIEMNTSVWEDQNEDESPKK----- 31
|||||
Qy 61 STAMKTCNCVAKTKLANGTSSMIVPKORKLSASYEKELCVKYFEQWSSDDQVEFVEHL 120
|||||
Db 32 -----SALWQISNGTSSVIVSRKRPSGNYQKEKOLCIKYFDQWSSDDQVEFVEHL 82
|||||
Qy 121 ISQCHYQHGHSINLYKPLMLQORDPITALPARGLDHIAENILSYLDAKSLCAAELVCKEY 180
|||||
Db 83 ISRCHYQHGHSINLYKPLMLQORDPITALPARGLDHIAENILSYLDAKSLCAAELVCKEWQ 142
|||||
Qy 181 RVTSDGMLWKKLIERMVRLTSLMRGLAERRGWGYLFKNKPPDGNAPPNSFYRALKPKII 240
|||||
Db 143 RVISEGMLWKKLIERMVRLTSLMRGLAERRGWGYLFKNKPPDGNAPPNSFYRALKPKII 200
|||||
Qy 241 QDIETIESNRCGRHSQRHCRSETSGKGYVCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300
|||||
Db 201 QDIETIESNRCGRHSQRHCRSETSGKGYVCLQYDDQKIVSGLRDNTIKIWDKNTLECK 260
|||||
Qy 301 RILTGHGTVLCLOYDERVITGSSDSTVRWVDMVTGEMNTLIHHCCEAVLHFRFNNGMM 360
|||||

```



```

255 GVTCLQYDDDKIISGLRDSIKIWKSSLECLVLTGHTGSLVLCLOYDERVIVTSSDST 314
329 VRVMDVNTGMLNTLIHHCBAVLHLRFNNGMMVTCSDRSIAVWDMASPTDITLRLVLVG 388
315 VRVMDVNTGVLNTLIHNEAVLHLRFNGLMTCSKDRSIAVWDMASATDITLRLVLVG 374
389 HRAAVNVDPDDKVIIVSASGDRITIKVNTSTCEVFTLNGHKGIAACLOYRDLVVSGSS 448
375 HRAAVNVDPDDKVIIVSASGDRITIKVNTSTCEVFTLNGHKGIAACLOYRDLVVSGSS 434
449 DNTIRLWDIECGACLRVLEGHEBELVRCIRFDNKRIVSGAYDGIKWMDLVAALDPRAPAG 508
435 DNTIRLWDIECGACLRVLEGHEBELVRCIRFDNKRIVSGAYDGIKWMDLVAALDPRAPAS 494
509 TGLCLRTLVHSGRVFRLOQDFEFOIVSSSHDDTILWDFNDPAAQAEPFRSPRTTYIS 568
495 TGLCLRTLVHSGRVFRLOQDFEFOIVSSSHDDTILWDFNVPPSAQNSTRSPRTTYIS 554
569 R 569
555 R 555

23HO PRELIMINARY; PRT; 563 AA.
-DEC-2001 (TrEMBLrel. 19, Created)
-MAR-2003 (TrEMBLrel. 23, Last sequence update)
box/WD40 repeat-containing protein HOS.
XXWL.
IS musculus (Mouse)
Karyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
mmalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
} _TaxID=10090;
RAIN=FVB/N;
atia N., Herter J.R., Slaga T.J., Fuchs S.Y., Spiegelman V.S.;
cuse homolog of HOS (mHOS) is overexpressed in skin tumors and
licated in constitutive activation of NF-kappaB;
mitted (JUN-2001) to the ENBL/GenBank/DBJ databases.
BL; AY038079; AAK72095.1; -.
D; MGI:2144023; Fbxwlb.
terPro; IPR001810; F-box.
am; PF00646; F-box; 1.
am; PF00400; WD40; 7.
INTS; PR00320; GPROTEINRPT.
ODOM; PD000018; WD40; 4.
IART; SM00256; FBOX; 1.
IART; SM00320; WD40; 7.
OSITE; PS50181; FBOX; 1.
OSITE; PS00678; WD REPEATS_1; 5.
OSITE; PS50082; WD REPEATS_2; 7.
OSITE; PS50294; WD REPEATS_REGION; 1.
beat; WD repeat.
QUENCE 563 AA; 64741 MW; 9AB562F3FF5E3496 CRC64;

Match 75.3%; Score 2286; DB 11; Length 563;
Local Similarity 75.7%; Pred. No. 1.2e-187;
es 440; Conservative 46; Mismatches 65; Indels 30; Gaps 7;

1 MDPAAVLQEKALFPMSSERE---DQNGGPPRKIIPEKNSLRQTNSCARCLNQ-E 55
1 MEP-DSVLTETIEMCSVPRLMIGCAN-----LVSMALSCIQMPS 44

56 TVCLASTA-MKTEN-----CVAKTKLANGTSSMIVPKQKLSAYEKEKELCVKYPQW 108
45 VRCLQNTSMEDQNEDESPKKSALWQISNGTSSVIVSKRPSEGNQKENDLCIKYFDQW 104

109 SESQVEFVHLLISOMCHYOHGHINSYLKPMQLQDFITALPARGLDHIAENILSYLDAKS 168
105 SESQVEFVHLLISRMCHYOHGHINSYLKPMQLQDFITALPEQGLDHAENILSYLDARS 164
169 LCAAEVCKEWMYRVTSQDMLWKKLIERMVTRTDSLNRGLAERRRGWQYLFKNKPPDGNAPP 228
165 LCAAEVCKEWMQVISEGMLWKKLIERMVTRTDLPLWKGLSERRGWDQYLFKNRPTDG--PP 222
229 NSFVRALYPKIIODIETIESNWECCGRHSQRHICRSETSGVYVCLOYDDQKIVSGLRDNT 288
223 NSFVRSYLPKIIODIETIESNWECCGRHNLQRIOCRSENSGVYVCLOYDDQKIISGLRDS 282
289 IKIWDKNTLBECKRILTGTGTVLCLQYDERVITGSSDSTVRVWDMVTGEMNTLIHHC 348
283 IKIWDKSSLECLVLTGHTGSLVLCLOYDERVITGSSDSTVRVWDMVTGELVNTLIHNE 342
349 AVLHLRFNNGMMVTCSDRSIAVWDMASPTDITLRLVLVGHRAAVNVDPDDKVIIVSASG 408
343 AVLHLRFNGLMTCSKDRSIAVWDMAFCHRYHFTPCSGWPRAAVNVDPDDKVIIVSASG 402
409 DRTIKVNTSTCEVFTLNGHKGIAACLOYRDLVVSGSSDNTIRLWDIECGACLRVLEG 468
403 DRTIKVNTSTCEVFTLNGHKGIAACLOYRDLVVSGSSDNTIRLWDIECGACLRVLEG 462
469 HEELVRCIRFDNKRIVSGAYDGIKWMDLVAALDPRAPAGTCLCLRTLVHSGRVFRLOQD 528
463 HEELVRCIRFDNKRIVSGAYDGIKWMDLVAALDPRAPASTCLCLRTLVHSGRVFRLOQD 522
529 EFOIVSSSHDDTILWDFNDPAAQAEPFRSPRTTYISR 569
523 EFOIVSSSHDDTILWDFNVPPSAQNSTRSPRTTYISR 563

RESULT 8
O44382 PRELIMINARY; PRT; 510 AA.
ID C044382;
AC C044382;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE SLIMB.
GN SLMB OR SLIMB OR CG3412.
OS Drosophila melanogaster. (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1];
RP SEQUENCE FROM N.A.
RX MEDLINE=98121115; PubMed=9461217;
RA Jiang J., Struhl G.;
RT "Regulation of the Hedgehog and Wntless signalling pathways by the F-
RT box/WD40-repeat protein Slmb.";
RL Nature 391:493-496(1998).
DR EMBL, AF032878; AAC38852.1; -.
DR FlyBase; FBgn0023423; slmb.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINRPT.
DR ProDom; PD000018; WD40; 4.
DR SMART; SM00256; FBOX; 1.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS00678; WD REPEATS_1; 5.
DR PROSITE; PS50082; WD REPEATS_2; 7.
DR PROSITE; PS50294; WD REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 510 AA; 58952 MW; DB0243D3730A5E8 CRC64;

Query Match 68.9%; Score 2089.5; DB 5; Length 510;
Best Local Similarity 78.3%; Pred. No. 7.7e-171;
Matches 394; Conservative 48; Mismatches 54; Indels 7; Gaps 4;

```



```

64 MKTENCVAKTCLANG---TSSMIVPKQKLSAS--YEKEKELCVKYFEQWSESQDQVEFVE 118
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinet K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary, and Imaginal disks;
RX MEDLINE=20245299; PubMed=10781936;
RA Miletich I., Limbourg-Bouchon B.;
RT "Drosophila null limb clones transiently deregulate Hedgehog-
independent transcription of wingless in all limb discs, and induce
decapentaplegic transcription linked to imaginal disc regeneration."
RL Mech. Dev. 93:15-26(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003733; AAF55853.1; -.
DR EMBL; AF222924; AAF63214.1; -.
DR EMBL; AF222923; AAF63213.1; -.
DR EMBL; AY118898; AAM50758.1; -.
DR FlyBase; FB00032423; slmb.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINBRPT.
DR PRODOM; PD000018; WD40; 4.
DR PROSITE; PS0181; FBOX; 1.
DR PROSITE; PS0678; WD_REPEATS_1; 5.
DR PROSITE; PS0082; WD_REPEATS_2; 7.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 510 AA; 58952 MW; F4D5DF126F58A012 CRC64;

Query Match 68.9%; Score 2089.5; DB 5; Length 510;
Best Local Similarity 78.3%; Pred. No. 7.7e-171;
Matches 394; Conservative 49; Mismatches 53; Indels 7; Gaps 4;

QY 64 MKTENCVAKTCLANG---TSSMIVPKQKLSAS--YEKEKELCVKYFEQWSESQDQVEFVE 118
4 METDKINDETN-SNAQAFITITMLYDPVRKXDSPTTYOTERELCFQYFTQWSESQDQVE 62
Db 4 METDKINDETN-SNAQAFITITMLYDPVRKXDSPTTYOTERELCFQYFTQWSESQDQVE 62
QY 119 HLISQCHYQHGHINSYKPMQLQDFITALPARGLDHIAENILSYLDAKSLCAELVCKE 178
63 HLLSRMCHYQHGHINSYKPMQLQDFITALPARGLDHIAENILSYLDAKSLCAELVCKE 122
QY 179 WYRTSDGMLWKKLIERWRTDSLWRGLAERRGWGLYLFKNKPPDGNAPNSFYRLPYK 238
123 WLRVISEGMLWKKLIERWRTDSLWRGLAERRGWGLYLFKNKPPDGNAPNSFYRLPYK 181
Db 123 WLRVISEGMLWKKLIERWRTDSLWRGLAERRGWGLYLFKNKPPDGNAPNSFYRLPYK 181

```

[illegible]

mus musculus (mouse);
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
NCBI_TaxID=10090;
(1)
RN
RP
SEQUENCE FROM N.A.
RA
Strausberg R.;
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; SC008552; AA08952.1; -
DR MGD; MGI:2144023; Fbxwlb.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 6.
DR PRINTS; PS00320; GP0TEINBRPT.
DR ProDom; PD000018; WD40; 3.
DR SMART; SM00320; WD40; 5.
DR PROSITE; PS00678; WD REPEATS 1; 5.
DR PROSITE; PS50082; WD REPEATS 2; 6.
DR PROSITE; PS50294; WD REPEATS REGION; 1.

FT	NON TER	1	SEQUENCE	252 AA	28424 MW	F71737CBD7A9F75F	CRC64	
			Query Match	40.5%	Score 1230	DB 11	Length 252	
			Best Local Similarity	92.9%	Pred. No. 1.9e-97			
			Matches 234	Conservative 7	Mismatches 11	Indels 0	Gaps 0	
Qy			318	RVIITGGSDSTVRVMDYNTGEMNTLIHCEAVLHLRFNNGMMVTCCKDRGIAVWDNASP	377			
Db			1	RVIITGGSDSTVRVMDYNTGEVNTLIHNEAVLHLRFNGLMVTCSKDRGIAVWDNAAA	60			
Qy			378	TDITLRRVLVGHRAAVNVDFDDKYIVSASGDRTIKWNTSTCEEVRTLNGHKGIACLQ	437			
Db			61	TDITLRRVLVGHRAAVNVDFDDKYIVSASGDRTIKWNTSTCEEVRTLNGHKGIACLQ	120			

Db	121	YRDLRVSGSDNTIRLWDIECGACURVLEGHEELVRCIRFDNKEIVSGAYDGKIKVWDL	180
Qy	498	VAALDPRAPATCLCLRTLVEHSGRVFRLQDFDFQIVSSSHDDTILIWDFLNDPAAQAEP	557
Db	181	QAALDPRAPATCLCLRTLVEHSGRVFRLQDFDFQIISSSHDDTILIWDFLNVPPSAQNET	240
Qy	558	RSPSTTYTYSR	569
Db	241	RSPSTTYTYSR	252

RESULT 12

Yoshikawa T., Nagai K., Sugano S., Aotsuka S., Aotsuka S., Yoshikawa Y.,
Tsunawara H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
Kakumura Y., Nagahara K., Masuho Y., Sasaki N.,
NEDO human cDNA sequencing project.;
Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
[BL; AK019133; BAA91986.1; -
InterPro; IPR001810; F-box;
fam; PF00646; F-box; 1.
roDom; PD000018; WD40; 2.
VART; SM00256; FBOX; 1.
VART; SM00320; WD40; 7.
ROSITE; PS50181; FBOX; 1.
ROSITE; PS00678; WD REPEATS 1; 5.
ROSITE; PS50082; WD REPEATS 2; 7.
ROSITE; PS50294; WD REPEATS_REGION; 1.
Protein; PS50294; WD REPEATS_REGION; 1.
Protein; PS50294; WD REPEATS_REGION; 1.
Repeat; WD repeat.
SEQUENCE 553 AA; 6280 MW; CA829C21986A3P2 CRC64;
Match 21.1%; Score 640; DB 4; Length 553;
Local Similarity 30.8%; Pred. No. 3e-46;
Matches 92; Conservative 92; Mismatches 211; Indels 72; Gaps 15;
14 KFMSSREDNNGEPPRKIIPEKNSLRQTYNSCARCLNQETVCLASTAMKTCNVAKT 73
17 KLDHGSEVRSFSGKPKCV-----SEYTTTGL-----VPCSA-----TPTTFGDL 58
74 KLANGTSSMIVPKQKLSASVEKEKEL--CVKYFEOWSESQDVEFVEHLISQMCHYQGH 131
59 RAANGQG-----QRRRITSVQPTGLQEWLKMFSQSGPEKLLADELIDSCPTQVKH 113
132 INSYLKPMQLORDITAPGLDHIENILSYLDAKSLCAELVCKEYRVTSQGLMVK 191
114 MQQVIEPQFQDFISLLP-----KELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWRE 169
192 LIERVVRTDSLWRGLAS-----BRGWQYLFPKNKPPDGNAPPNSFYRALPKIIQDIETI 246
170 KCKE-----EGIDEPHLHKRK-----VTKPGFIHSPKSA-----IRQ--HRI 207
247 ESNWRCGRHSQRIHCRSETSKGYVCLQYDDQKIVSGLRDNTIKINDKNTLECKRIILTG 306
208 DTNWRGELSKPKV-LKGGDDHVTICLQFCGNRIVSGDDNTLKWSAVTGKCLRTLUGH 266
307 TGSVLCLOYDERVITGSSDSTVRVWDVNTGEMNLTIHCEAVLHLRFNNGMMVTCSD 366
267 TGGVWSSQMRDNIISGSTDRTLKVMNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRD 326
367 RSIADVMASTPTDITLRRVLVGHRAAVNVDPDKIVSASGDRTIKVMNTSTCFVRTL 426
327 ATLVRWDIETGQCL---HVLMGHVAARVCQYDGRVVSAGYDFWVKVWDPTETCLHTL 383
427 NGHKGRIACQYDRDLRVVSGSSDNTIRLWDIEGACLRVLEGEELVRCIRFONKRVISG 486
384 QGHTRVYSLQFDGIHVVSGLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSG 443
487 AYDGKIKVLDVAALDPAPAGTLCRLTV---EHSGRVFRLOFDFQIVSSSHDDTILI 543
444 NADSTVKIWDIKTG-----QCLQTLQGNPKHQSAVTCLOFNKNFVITSSDDGTVKL 494
544 WD 545
495 WD 496

14

6R12 PRELIMINARY; PRT; 561 AA.

-DEC-2001 (TrEMBLrel. 19, Created)

-DEC-2001 (TrEMBLrel. 19, Last sequence update)

-MAR-2003 (TrEMBLrel. 23, Last annotation update)

box protein FBX30 (Fragment).

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1] SEQUENCE FROM N.A.
RP MEDLINE=20003061; PubMed=10531037;
RX Winston J.T., Koepf D.M., Zhu C., Ellledge S.J.;
RT "A family of mammalian F-box proteins.";
RL Curt. Biol. 9:1180-1182 (1999).
RN [2] SEQUENCE FROM N.A.
RA Koepf D.M., Winston J.T., Harper J.W., Ellledge S.J.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RX EMBL; AF383178; AAK60269.1; -
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINBRPT.
DR ProDom; PD000018; WD40; 2.
DR SMART; SM00256; FBOX; 1.
DR SMART; SM00320; WD40; 8.
DR PROSITE; PS50181; FBOX; 1.
DR PROSITE; PS00678; WD REPEATS 1; 5.
DR PROSITE; PS50082; WD REPEATS 2; 7.
DR PROSITE; PS50294; WD REPEATS_REGION; 1.
KW Repeat; WD repeat.
FT NON TER 1
SQ SEQUENCE 561 AA; 63165 MW; B81CCLB2206B0D88 CRC64;
Query Match 21.1%; Score 640; DB 4; Length 561;
Best Local Similarity 30.8%; Pred. No. 3.1e-46;
Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;
QY 14 KFMSSREDNNGEPPRKIIPEKNSLRQTYNSCARCLNQETVCLASTAMKTCNVAKT 73
DB 25 KLDHGSEVRSFSGKPKCV-----SEYTTTGL-----VPCSA-----TPTTFGDL 66
QY 74 KLANGTSSMIVPKQKLSASVEKEKEL--CVKYFEOWSESQDVEFVEHLISQMCHYQGH 131
DB 67 RAANGQG-----QRRRITSVQPTGLQEWLKMFSQSGPEKLLADELIDSCPTQVKH 121
QY 132 INSYLKPMQLORDITAPGLDHIENILSYLDAKSLCAELVCKEYRVTSQGLMVK 191
DB 122 MQQVIEPQFQDFISLLP-----KELALYVLSFLEPKDLLQAAQTCRYWRILAEDNLLWRE 177
QY 192 LIERVVRTDSLWRGLAS-----BRGWQYLFPKNKPPDGNAPPNSFYRALPKIIQDIETI 246
DB 178 KCKE-----EGIDEPHLHKRK-----VTKPGFIHSPKSA-----IRQ--HRI 215
QY 247 ESNWRCGRHSQRIHCRSETSKGYVCLQYDDQKIVSGLRDNTIKINDKNTLECKRIILTG 306
DB 216 DTNWRGELSKPKV-LKGGDDHVTICLQFCGNRIVSGDDNTLKWSAVTGKCLRTLUGH 274
QY 307 TGSVLCLOYDERVITGSSDSTVRVWDVNTGEMNLTIHCEAVLHLRFNNGMMVTCSD 366
DB 275 TGGVWSSQMRDNIISGSTDRTLKVMNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRD 334
QY 367 RSIADVMASTPTDITLRRVLVGHRAAVNVDPDKIVSASGDRTIKVMNTSTCFVRTL 426
DB 335 ATLVRWDIETGQCL---HVLMGHVAARVCQYDGRVVSAGYDFWVKVWDPTETCLHTL 391
QY 427 NGHKGRIACQYDRDLRVVSGSSDNTIRLWDIEGACLRVLEGEELVRCIRFONKRVISG 486
DB 392 QGHTRVYSLQFDGIHVVSGLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSG 451
QY 487 AYDGKIKVLDVAALDPAPAGTLCRLTV---EHSGRVFRLOFDFQIVSSSHDDTILI 543
DB 452 NADSTVKIWDIKTG-----QCLQTLQGNPKHQSAVTCLOFNKNFVITSSDDGTVKL 502
QY 544 WD 545
||

503 WD 504

15

96LE0 PRELIMINARY; PRT; 589 AA.

1-DEC-2001 (TremBLrel. 19, Created)

1-DEC-2001 (TremBLrel. 19, Last sequence update)

1-MAR-2003 (TremBLrel. 23, Last annotation update)

-box protein SEL10.

EU10.

omo sapiens (Human).

lkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

ammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

CBI_TaxID=9606;

11

SEQUENCE FROM N.A.

1 J., Pauley A.M., Myers R.L., Shuang R., Brashler J.R., Yan R.,

hl A.E., Gurney M.E.;

SEL-10 Interacts with Presenilin 1, Facilitates Its Ubiquitination,

nd Alters A-beta Production.";

abmitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

MBL AY008274; AAG16640.1; ..

nterPro; IPR001810; F-box.

terPro; IPR001880; WD40.

fam; PF00646; F-box; 1.

fam; PF00400; WD40; 7.

RINTS; PR00320; GPROTEINRPT.

coDom; PD000018; WD40; 2.

4ART; SM00256; FBOX; 1.

4ART; SM00320; WD40; 8.

ROSITE; PS50181; FBOX; 1.

ROSITE; PS00678; WD_REPEATS_1; 5.

ROSITE; PS50082; WD_REPEATS_2; 7.

ROSITE; PS50294; WD_REPEATS_REGION; 1.

peat; WD repeat.

QUENCE 589 AA; 66120 MW; 2AFB6E8A36B6E8DE CRC64;

, Match 21.1%; Score 640; DB 4; Length 589;

Local Similarity 30.8%; Pred. No. 3.3e-46;

ies 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

14 KFMNSEREDCNNGEPKPIPEKNSLPQTYNSCARLCLNQETVCLASTAMKIENCVAKT 73

53 KLDHGEVRSFSLGKKPKV-----SEYTSITGL-----VPCSA-----TPTFGDL 94

74 KLANGTSSMIVPKQKLSAYEKEKEL--CVKYPEQWSESQVFEVHLISQMCYQGH 131

95 RAANGOG-----QORRITSVQPTGLQEWLKMFSQSGPEKLLALDELIDSCPTQVKH 149

132 INSYLKPMLORFITAPGLDHTAENILSYLDAKSLCAELVCKEYRVYVTSQGLWKK 191

150 MMQVIEPQFQDFISLLP-----KELALVLSFLEPKDLQAAQTCRYVRILAEDNLLWRE 205

192 LIERVMYRDSLWRGLAE-----RRGWGYLFKNKPPDGNAPPNSFYRALYPIQDIETI 246

206 KCCE-----EGIDEPLHIKRRK-----VIKEGFIHSPKSAI-----IRQ--HRI 243

247 ESNWRCGRHSIORHCRSETSGVYCLQYDQKIVSGLRDNTIKIWKNTLECKRIITGH 306

244 DTNWRREGELKSPKV-LKGHDHVTITCLQFCGNRIVSGDDNTLKVMSAVITGKCLRTLVGH 302

307 TGSVLCLQYDERVITGSDSTVRYVDVNTGEMTLIHCEAVLHLRFNNGMMVTCSD 366

303 TCGWSSQNRDNIISGSDRTLKVNNAETGECIHTLYGRTSVRCMHLHEKRVVSGSRD 362

367 RSIAYNDASPTDITLRLRVLGHRAAVNVDPDDRYIVSASGDRITKIVNTSTCEFVRTL 426

363 ATLRYWDIETGQCL--HVLGMHVAARVCQVQDGRVVSQVDFWVWVDPETETCLHTL 419

427 NGHKRGIACLQYRDLVTSQSDNTIRLWDIECGACLRVLEGHBEELVRCIFPDNKRIVSG 486

Db 420 QCHTRVYSLQFDGIHVYVSGSLDTSIRVWDVETGNCIHTLTGHQSLTSGMELKDNILVSG 479

Qy 487 AYDGKIKVWDLVAALDPRAPAGTLCRLTLV---EHSGRVPRLOQDFEQIVSSSHDDTILI 543

Db 480 NADSTVKIWDIKTG-----OCLQTLQGPKNKHOSAVTCLQFNKNFVITSSDDGTVKL 530

Qy 544 WD 545

Db 531 WD 532

Search completed: October 22, 2003, 10:25:14

Job time : 110 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

ein - nucleic search, using frame_plus_p2n model
October 22, 2003, 13:59:32 : Search time 4994 Seconds
(without alignments)
4661.109 Million cell updates/sec

US-09-601-168B-2
score: 3034
e: 1 MDPAAVLQKALKFMSSE.....PAAQAEPPSPSRVTYIYSR 569

table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

d: 2888711 seqs, 20454813386 residues
umber of hits satisfying chosen parameters: 5777422

DB seq length: 0
DB seq length: 2000000000

ocessing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

line parameters:
frameat_p2n.model -DSV=xlp
2_1/USPTO.spool_p/US09601168/runat 22102003 100546 17778/app.query.fasta_1.711
Embl -CPWT=fastab -SUPFIX=rg -MINMATCH=0.1 -LOOSCL=0 -LOOEXT=0
bits -START=1 -END=-1 -MATRIX=BIOSUM62 -TRANS=human40.cdi -LIST=45
GN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
S09601168 @CGN 1 1 3745 @runat 22102003 100546 17778 -NCPU=6 -ICPU=3
P -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
MEOUT=120 -WARN_TIMECUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
T=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

s :
GenEmbl:*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vi:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pin:*
35: em.htg.rod:*
36: em.htg.nam:*
37: em.htg.vrt:*
38: em.sy:*
39: em.htgo.hum:*
40: em.htgo.mus:*
41: em.htgo.other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	3034	100.0	2151	6	AX019507 Sequence
2	3034	100.0	2151	6	AX057166 Sequence
3	3034	100.0	2151	6	BD131778 Human bet
4	3034	100.0	2151	9	AF129530 Homo sapi
5	3034	100.0	2151	9	Y14153 Homo sapien
6	3006	99.1	1818	9	AF101784 Homo sapi
7	3006	99.1	2308	9	BC027994 Homo sapi
8	2997	98.8	1712	10	AF099932 Mus muscu
9	2997	98.8	2081	10	BC003989 Mus muscu
10	2997	98.8	2175	10	AF081887 Mus muscu
11	2990	98.5	1710	10	AF112979 Mus muscu
12	2978	98.2	1979	10	AF110396 Mus muscu
13	2597	85.6	1671	5	M98268 African Cla
14	2400	79.1	2252	9	AB033280 Homo sapi
15	2400	79.1	4395	9	BC026213 Homo sapi
16	2384.5	78.6	1653	9	AF176022 Homo sapi
17	2384.5	78.6	2274	9	AB033281 Homo sapi
18	2384.5	78.6	4230	9	AB014596 Homo sapi
19	2380	78.4	4030	10	BC034261 Mus muscu
20	2361	77.8	2134	9	AB033279 Homo sapi
21	2360	77.8	3874	10	AB093260 Mus muscu
22	2330	76.8	1970	10	AY038079 Mus muscu
23	2250	74.2	2530	5	BC045356 Danio rer
24	2129.5	70.2	2514	3	AK114399 Ciona int
25	2128	70.1	2445	3	AB076893 Ciona int
26	2089.5	68.9	2154	3	AF032878 Drosophil
27	2089.5	68.9	2367	3	AF222924 Drosophil
28	2089.5	68.9	2534	3	AY118898 Drosophil
29	2089.5	68.9	3532	3	AF222923 Drosophil
30	1908.5	62.9	67008	9	AF000252 Homo sapi
31	1908.5	62.9	100000	9	AP000212 Homo sapi
32	1908.5	62.9	149298	9	AP000031 Homo sapi
33	1908.5	62.9	340000	9	AP001711 Homo sapi
34	1685	55.5	173634	3	AC009344 Drosophil
35	1685	55.5	235707	3	AE003733 Drosophil
36	1683.5	55.5	292639	2	AC112602 Rattus no
37	1653.5	54.5	40472	2	AC014085 Drosophil
38	1638.5	54.0	2700	3	AF275253 Caenorhab
39	1503.5	49.6	17344	3	U28730 Caenorhabdi
40	1498	49.4	1443	3	AF339101 Heteroder
41	1230	40.5	2983	10	BC008552 Mus muscu
42	1189	39.2	1827	5	U63922 Xenopus lae
43	1180	38.9	934	5	XLU63921 Xenopus lae
44	1163	38.3	657	6	AX505273 Sequence
45	1075	35.4	1217	3	AK115763 Ciona int

ALIGNMENTS

```

7
ION AX019507 2151 bp DNA linear PAT 07-SEP-2000
ON Sequence 1 from Patent WO9938969.
AX019507
AX019507.1 GI:10043427
S .
synthetic construct
ISM synthetic construct
artificial sequences.
CE 1
RS Arenzana,S.F., Concordet,J.P., Kroll,M., Durand,H., Benarous,R. and
Margottin,F.
Protein humaine beta-trop.
AL Patent: WO 9938969-A1 05-AUG-1999;
ARENZANA SEISDEDOS FERNANDO (FR); CONCORDET JEAN PAUL (FR); INST
NAT SANTE RECH MED (FR); KROLL MATHIAS (FR); DURAND HERVE (FR);
BENAROUS RICHARD (FR); MARGOTTIN FLORENCE (FR); PASTEUR INSTITUT
(FR)
S Location/Qualifiers
urce 1. 2151
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/ncfe="ADNC codant pour la proteine humaine BTRCP"
70..1779
/notes="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAC07740.1"
/db_xref="GI:10043428"
/translation="NDPAVLQEKALKFMNSSEREDNCNCEPPPKIIPKNSLRQTV
NSCARCLNQEVCLASTAWKTECNVAKTKLANGTSSMLVPKQKLSASVKEKELCV
KFEQSESDQEFVHLLISQCHVQGHINSYLKPLMQRFITLALPARGLDHIAENI
LSYLDASLCAELVELCKEYRVTSQGLMWRVDSLRGLABRGGWGLFK
NKPDDNAPNPFYALYKIIQDIETIESNRWCRHSLORIHCRSETSKGVYCLQYD
DQIVSGLRNTIKIWDKNTLBEKRLITLGTGSLVLCLOYDERVITGSSDSFVRWQV
NTGEMLNTLIHHCCEAVLHLRFNMVMVTCSDRSIAVMDMASPTDITLRRVLVGHRAA
VNVVDKDYIVSASGDRITIKYNTSTCFEFTLNGHKGIACLQYRDLVVSQSDN
TILWLDIEGACLRVLECHEELVRCIRFPNKRIVSGAYDGGIKIWDWLVAAALDPRAPAG
TLCLRILVSHSGVRFRLQDFQDFQISSHDDTILILWDFLNDPAQAEPSPSRITY
ISR"
UNT 628 a 467 c 513 g 543 t

nt Scores:
c.: 3.75e-274 Length: 2151
Similarity: 3034.00 Matches: 569
cal Similarity: 100.00% Conservative: 0
atc: 100.00% Mismatches: 0
6 Indels: 0
Gaps: 0

01-168B-2 (1-569) x AX019507 (1-2151)
1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerGlu 20
70 ATGCACCCGCCGAGGCGGTCTGCAAGAGAGAGCACTCAAGTTTATGATTAATCTCTCAGAG 129
21 ArgGluAspCysAsnAndlyGluProProArgLysLilleProGluLysAsnSerLeu 40
130 AGAGAAGACTGTAAATATGCGCAACCCCTTAGAAGATAATACAGAGAAGAAATTCACCT 189
41 ArgGlnThrTyrrAsnSerCysAlaArgLeuLysLeuAsnGlnGluThrValCysLeuAla 60
190 AGACAGACATACACACAGCTGCGCAGACTCTGCTTAACCAAGAACACATGATGTTAGCA 249
61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
250 AGCACTCTATGAGACTGAGAAATTTGTGGCCCAAAACAAACTTGCCCAATGGCACTTCC 309
81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrrGluLysGluLysGlu 100
310 AGTATGATTTGTGCCCAAGCAACGGAACCTCTAGCAGGCTATGAAAGAAAGAAAGCACTG 369

```

461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
 1450 GCATGTTTACAGTGTAGAGGCCATGAGGAATTGGTGGTTGTATTCGATTGATAAC 1509
 481 LysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuAlaAla 500
 1510 AAGAGGATAGTCAGTGGGGCTATGATGGAAAAATTAAGATGGGGATCTTGTGGTGTCT 1569
 501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
 1570 TTGGACCCCGTCTCTCTGAGGACACATCTGTCTACGGACCCCTTGTGGAGCATTCGGGA 1629
 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
 1630 AGAGTTTTTCGATACAGTTTGTATGATTCAGATTGTCTAGTTCACATGATGACACA 1689
 541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProArgSerPro 560
 1690 ATCTCATCTGGGACTTCTTAATGATCCAGTGGCCCAAGCTGAACCCCGGTTCCCT 1749
 561 SerArgThrTyrThrIleSerArg 569
 1750 TCTCGAACATACACCTACATCTCCAGA 1776

2
 6 AX057166 2151 bp DNA linear PAT 17-JAN-2001
 ION Sequence 42 from Patent WO0075184.
 ON AX057166
 AX057166.1 GI:12309979

S Homo sapiens (human)
 ISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CE 1
 RS Zhang, H., Tsvetkov, L.M. and Kondo, T.
 AL Modulation of protein levels using the scf complex
 Patent: WO 0075184-A 42/14-DEC-2000;
 YALE UNIVERSITY (US)
 S Location/Qualifiers
 1..2151
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

UNT 628 a 467 c 513 g 543 t

nt Scores:
 O.: 3.75e-274 Length: 2151
 3034.00 Matches: 569
 Similarity: 100.00% Conservative: 0
 cal Similarity: 100.00% Mismatches: 0
 Indels: 0
 Gaps: 0

01-168B-2 (1-569) x AX057166 (1-2151)

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
 70 ATGGACCCGCCGAGGGGTGCTGCAAGAGAGAGGCACTCAAGTTTATGAATTCCTCAGAG 129
 21 ArgGluAspCysAsnAsnGlyGluProProArgLysIleLeuProGluLysAsnSerLeu 40
 130 AGAGAAGACTGTAATATGCGCAACCCCTTAGAAGATAATACAGAGAAGAAATTCACCT 189
 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
 190 AGACAGACATACACAGCTGTGCGCAGACTCTGCTTAAACCAAGAAACAGATGTTTAGCA 249
 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
 250 AGCACTGCTATGAAGACTGAGATTGTGTGGCCCAACAAACAACTTGCCCAATGGCACTTCC 309

QY 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
 Db 310 AGTATGATTGTGCCAAGCAACGAAACTCTCAGCAGACTATGAAAGGAAAAGGAACTG 369
 QY 101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
 Db 370 TGTGTCAATACTTTGAGCAGTGGTCAGAGTCAGATCAAGTGGAAATTTGTGGAACATCTT 429
 QY 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
 Db 430 ATAATCCCAATGTCTATTACCAACATGGGCAATAAACTCGTATCTTAAACCTTGTGTG 489
 QY 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
 Db 490 CAGAGAGATTTCATAACTGTCTGCCAGCTCGGGATTTGGATCATATCCGTGAGAACATT 549
 QY 161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 180
 Db 550 CTGTACATCTGGATGCAAAATCACTATGTCTGCTGCTGAACTTGTGTGCAAGGAATGGTAC 609
 QY 181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
 Db 610 CGAGTGACCTCTGATGGCATGCTGTGGAAGAACTTATCGAGAGAAATGGTCAGGACAGAT 669
 QY 201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220
 Db 670 TCTCTGTGGAGAGGCGCTGGCAGAACGAAGAGGATGGGACAGTATTATTTCAAAAACAAA 729
 QY 221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
 Db 730 CCTCCTGACGGGAATGCTCTCCCAACTCTTTTATAGAGCACTTTATCTTAAAAATTATA 789
 QY 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
 Db 790 CAAGACATTGAGACAATAGAACTCTAATTGGAGATGTGGAGACATAGTTTACAGAAATT 849
 QY 261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
 Db 850 CACTGCCGAAGTGAACCAAGCAAAAGGAGTTTACTGTTTACAGTATGATCAGAAAAATA 909
 QY 281 ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300
 Db 910 GTAAGCGCGCTTCGAGACAAACAAATCAAGATCTGGGATAAAAAACACATTTGGAAATCAAG 969
 QY 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
 Db 970 CGAATTCACAGGCCCATACAGTTTCAGTCTCTCTCTCAGTATGATGAGAGAGTGATC 1029
 QY 321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
 Db 1030 ATACAGGATCATCGGATTCACGGTCAGAGTGTGGATGTAATAACAGGTGAATTCCTA 1089
 QY 341 AsnThrLeuIleHisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
 Db 1090 AACAGTTGATTACCAATTGTGAAGCAGTTCTGCATTTGCTTCAATATGCGCATGATG 1149
 QY 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
 Db 1150 GTGACCTGTCTCCAAAGATGTTCCATTGTGTGTGGATATGCGCTCCCACTGACATT 1209
 QY 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
 Db 1210 ACCCTCCGAGGGTGTGGTCCGACACCGAGCTGCTCTCAATGTTCTGAGACTTTGATGAC 1269
 QY 401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
 Db 1270 AAGTACATTGTTTCTGCATCTGGGGATAGAACTATATAAGGTATGGAACACAAAGTACTTGT 1329
 QY 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
 Db 1330 GAATTTGTAGGACCTTAATATGACACAAACGAGGCAATGCGCTGTTTGCAGTACAGGAC 1389

441	ArgLeuValValSerGlySerSerAspasnThrIleArgLeuTrpAspIleGluCysGly	460
1390	AGGCTGTAGTGTAGTGGCTCATCTCACAACTATCATGATTATGGACATGAATGTGGT	1449
461	AlaCysLeuArgValLeuGluGlyHisIscLgLuLeuValArgCysIleArgPheAsn	480
1450	GCATGTTTAGCAAGTGTAAGGCCATCGAATGGTCGTTGATTTCGATTGATAAC	1509
481	LysArgIleValSerGlyAlaTrpAspGlyLysIleLysValTrpAspLeuValala	500
1510	AAGAGGATAGTCACTGGGCGCTTATCATGCAAAAATTAAAGTGTGGGATCTTGTGGCTGT	1569
1570	TTGACCCCGTCTCTCGAGGACACTGTCTACGGACCCTTGTGAGCATTCGGGA	1629
521	ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspThr	540
1630	AGAGTTTTTCGACTCAGTTTCATGAATTCAGATTGTCACTAGTAGITCATCATGACACA	1689
541	IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro	560
1690	ATCTCATCTGGGACTTCTAAATGATCCAAGCTGCCAAGCTGAACCCCGGTTCCTCT	1749
561	SerArgThrTrpThrTrpIleSerArg	569
1750	TCTCGAACATACACCTTACATCTCCAGA	1776
B	BD131778	2151 bp DNA linear PAT 18-SEP-2002
ION	Human beta TrCP protein.	
ON	BD131778	
S	BD131778.1 GI:23226723	
SM	JP 2002501746-A/1.	
SE	Synthetic construct	
SS	artificial construct	
AL	artificial sequences.	
	1 (bases 1 to 2151)	
	Benarous, R., Margottin, F., Durand, H., Seisdedos, F.A., Kroll, M. and	
	Concordet, J.P.	
	Human beta TrCP protein	
	Patent: JP 2002501746-A 1 st 22-JAN-2002;	
	INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE, INSTITUT	
	PASTEUR	
	OS Artificial Sequence	
	PN JP 2002501746-A/1	
	PD 22-JAN-2002	
	PF 29-JAN-1999 JP 2000529429	
	PR 30-JAN-1998 FR 98/01100, 09-DEC-1998 FR 98/15545 PI	
	RICHARD BENAROUS, FLORENCE MARGOTTIN, HERVE DURAND, PI FERNANDO	
	ARENZANA SEISDEDOS, MATTHIAS KROLL, JEAN PAUL CONCORDET PC	
	CI2N15/00, A01K67/027, A61K38/00, A61K45/00, A61P29/00, A61P31/12, PC	
	A61P35/00,	
	PC	
	C07K14/47, CI2N1/15, CI2N1/19, CI2N1/21, CI2N5/10, CI2Q1/68/CI2P21/ PC	
	02,	
	CC CI2N15/00, A61K37/02, CI2N5/00	
	CC Description of the artificial sequence :	
	ADNC coding for human	
	CC (TrCP	
	CC protein	
	FH key Location/Qualifiers	
	FT CDS (70)..(1776).	
	Location/Qualifiers	
	1..2151	
	/organism="synthetic construct"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:32630"	
DNT	, 628 a 467 c 513 g 543 t	
	Scores:	

Pred. No.:	3,75e-274	Length:	2151
Score:	3034.00	Matches:	569
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-601-168B-2 (1-569) x BD131778 (1-2151)

QY	1	MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu	20
DB	70	ATGACCGCGCGAGCGGTCTGCTCAGAGAGGCACTCAAGTTTATGAATTCCTCAGAG	129
QY	21	ArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLysAsnSerLeu	40
DB	130	AGAGAAGACTGTAAATATGGCGAAACCCCTAGGAAGATAATACAGAGAAGAAATTCACAT	189
QY	41	ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla	50
DB	190	AGACAGACATACCAACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTATGTATTGAC	249
QY	61	SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer	80
DB	250	AGCACTGCTATGAGACTGACAAATGTGTGCCAAACAAACTTGCCAAATGCACTTCC	309
QY	81	SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu	100
DB	310	AGTATGATGTGGCCCAAGCAACGGAAACTCTCAGCAAGCTATGAAAGGAAAGGAAGCTG	369
QY	101	CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu	120
DB	370	TGTGTCAAACTACTTTGACAGTGTGTGAGCTCAGATCAAGTGGAAATTTGTGGNACAATCT	429
QY	121	IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu	140
DB	430	ATATCCCAAAATGTGTCAATTACCAACATGGGCACATAAACTCGTATCTTAAACCTATGTTG	489
QY	141	GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle	160
DB	490	CAGAGAGATTTCAATACTGCTCTCCAGCTCGGGATGAGTATCATATCGCTGAGAACATTT	549
QY	161	LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr	180
DB	550	CTGTGATACCTGGATGCCAATCACTATGTGCTGCTGAACTTGTGTGCAAGNAATGGTAC	609
QY	181	ArgValThrSerAspGlyMetLeuTrpLysLeuIleGluArgMetValArgThrAsp	200
DB	610	CGAGTACCTCTGATGGCATGCTGTGGAAGAAGCTTATCGAGAAGATGGTCAGACAGAT	669
QY	201	SerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLys	220
DB	670	TCTCTGTGAGAGCCTGGCAAGACGAAGAGATGGGACAGTATTTATTCAAAACAA	729
QY	221	ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle	240
DB	730	CCTCTCGCGGAATGCTCTCCCAACTCTTTTATAGACACTTTTATCCTAAATATTA	789
QY	241	GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle	260
DB	790	CAAGACATTTGAGACAAATAGATCTAAATTTGAGATGTGGAAGACATAGTTTACAGAGAAT	849
QY	261	HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle	280
DB	850	CACTGCCGAAGTGAAACAGCAAGAGAGTTTACTGTTTACAGTATGATGATCAGAAATA	909
QY	281	ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys	300
DB	910	GTAAGCGCCTTCGAGACAAACACATCAAGATCTGGGATAAAACACATTTGGATCCAG	969
QY	301	ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle	320
DB	970	CGAAATCTTCAGAGCCCATACAGGTTTCACTCTCTGCTCCAGTATGATGAGAGATGATC	1021

nt Scores:


```

141  GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
490  CAGAGAGATTTTCATTAACGTCTCGCAGCTCGGGATTGGATCATATCGCTGAGAACATT 549
161  LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysGlyGluTrpTyr 180
550  CTGTATACCTCGATGCCAAATCACTATGTGTGCTGTAACCTGTGTGCAAGGAATGGTAC 609
181  ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
610  CAGAGTACCTCTGATGGCATGCTGTGGGAAGCTTATCGAGAGAAATGCTCAGACAGAT 669
201  SerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220
670  TCTCTGTGAGAGGCGCTGCAGAACAGAGAGATGGGACAGTATTTATTCAAAAACAAA 729
221  ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
730  CTTCTTGACGGGAATGCTCTCCCACTCTTTTATAGAGCACTTTATCTTAAATATATA 789
241  GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
790  CAAGACATTGAGACAATAAGATCTAATGGAGATGTGAAGACATAGTTTACAGAGAATT 849
261  HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
850  CACTCGCGAAGTGAACAAGCAAGAGAGTTTACTGTTTACAGTATGATGATCAGAAAAA 909
281  ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300
910  GTAAGGGGCTTCGAGACAACAACATCAAGATCTGGGATATAAAACACATCGAATGCAAG 969
301  ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
970  CGAATTCTCACAGGCCATACAGGTCCTCTCTCTCCAGTATGATGATGAGAGATGATC 1029
321  IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
1030  ATAACAGGATCATCGGATTCACCGTCAAGTGTGGATGTAAATACAGGTGAATGCTA 1089
341  AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
1090  AACACGTTGATTCACCATTTGTAGCAGTTCTGCACCTTCGCTTCAATTAATGGCATGATG 1149
361  ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
1150  GTGACCTGCTCCAAAGATCGTTCCATTGCTGTATGGGATATGGCCTCCCCCACTGACATT 1209
381  ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
1210  ACCCTCCGAGGGTGTGTGCGGACACCGAGCTGTGTCAATGTTGTAGACTTTGATGAC 1269
401  LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
1270  AAGTACATTGTTTCTGCATCTGGGATAGACATAATAAGGTATGGAACACAGTACTTGT 1329
421  GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
1330  GAATTTGTAGGACCTTAATGGACAACAACAGGAGCATTCCTGTTTGGGTACAGGGAC 1389
441  ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
1390  AGGCTGGTAGTGCTGCTCACTGACACATATCAGATTATGGACATAGATGTGGT 1449
461  AlaCysLeuArgValLeuGluGlyHisGluLeuValArgCysIleArgPheAspAsn 480
1450  GCATGTTTACGAGTGTAGAGGCCATGAGGAATTTGGTGGTGTATTTCGATTTGATAAC 1509
481  LysArgIleValSerGlyValTrpAspGlyLysIleLysValTrpAspLeuValAlaAla 500
1510  AAGAGGATAGTCACTGGGCGCTATGATGAAAAAATTAAGTGTGGGATCTTGTGCTGCT 1569

```

501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
1570 TTGGACCCCGGTCTCTGCGAGGACACTGTCTACGACCTTGTGGAGCATTCGCGA 1629
521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
1630 AGAGTTTTTCGACTACAGTTTGAATTCAGATTTGTCAGTAGTTCATGATGACACA 1689
541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
1690 ATCCTCATCTGGACTTCTCTAATGATCCAGTGCAGCTGACACCCCGGTTCCCT 1749
561 SerArgThrTyrThrTyrIleSerArg 569
1750 TCTCGAACATACACCTACATCTCCAGA 1776

RESULT 5
HSBTRCP
LOCUS HSBTRCP 2151 bp mRNA linear PRI 07-FEB-2003
DEFINITION Homo sapiens mRNA for beta-transducin repeat containing protein.
ACCESSION Y14153
VERSION Y14153.1 GI:2995193
KEYWORDS beta-transducin repeats; beta-TRCP gene; WD repeat.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Margottin, F., Bour, S.P., Durand, H., Selig, L., Benichou, S., Richard, V., Thomas, D., Strebel, K. and Benarous, R.
TITLE A novel human WD protein, h-beta TrCP, that interacts with HIV-1 Vpu connects CD4 to the ER degradation pathway through an F-box motif.
JOURNAL Mol. Cell 1 (4), 565-574 (1998)
MEDLINE 98325370
PUBMED 9660940
REFERENCE 2 (bases 1 to 2151)
AUTHORS Benarous, R.
TITLE Direct Submission
JOURNAL Submitted (03-JUL-1997) R. Benarous, INSERM - I.C.G.M., Laboratoire Interactions Proteiques, CHU Cochin, 24 rue de Fg.St-Jacques, 75014 Paris, FRANCE

FEATURES
Location/Qualifiers
1..2151
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="Jurkat T-cell"
/cell_type="T lymphocyte"
/tissue_type="lymphoid"
1..2151
/gene="beta-TRCP"
70..1779
/gene="beta-TRCP"
/codon_start=1
/product="beta-transducin repeats containing protein"
/protein_id="CAA74572.1"
/db_xref="GI:2995194"
/db_xref="GOA:Q9Y297"
/db_xref="SWISS-PROT:Q9Y297"
/translation="MDPAEVLQEKALKFMNNSEREDCNNGEPKRIIPEKNSLRQTY
NSCARLCLNQETVCLASTAMKTCNVAKTNGKLANGTSMIVPKQKLASYEKEKELCV
KFEQWMSQDVEFVEHLISQWCHYOHGHINSYLPKMLQRFITALPARGDLHIAENI
LSYLDKSLCAELNCKEWRVTSQMLWKLLERWRTDSLRGLAERRGQVLFK
NPPGNAFPNSFYALYPIKIIQDIETIESNWRCSQRHSQRIHCRSETSGKYCLQYD
DKIVSGRDNTIKIDKNTLECKILGTGSLVLCQYDERVIITGSSDSTVRVDV
KTFEMLNTLIHHEAVLHFRNNGMVTCKSDRSIAVWDMASPTDITLRLVLRHRAA
VNWVPDDKIIVSAGSDRTIKVNTSTCFVTRTLNGHKRGIACIQYRDLRVVSGSDN
TIRLWDIEGACLRVLEGEELVRICIRFDNKRIVSGAYDGKIKWDLVAALDPRPAG
TLCLRTLVEHSRGRVRLQDFEQIVSSSHDDTILLIWDFLNDPAAQAEPSPSRITY
ISR"
259..544
/rpt_family="WD"

/rpt unit=516. .544
UNT 628 a 467 c 513 g 543 t

nt Scores:

3.75e-274 Length: 2151
3034.00 Matches: 569
Similarity: 100.00% Conservative: 0
cal Similarity: 100.00% Mismatches: 0
atch: 100.00% Indels: 0
9 Gaps: 0

01-168B-2 (1-569) x HSBTRCP (1-2151)

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
70 ATGGACCCCGCGCGCGTGTGCAAGAGAGAGGACCTCAAGTTTATGAATTCCTCAGAG 129
21 ArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLysAsnSerLeu 40
130 AGAGAGAGACTGTAATAATGGCGAACCCCTAGAGAGATAATACAGAGAGAAGATTCATT 189
41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
190 AGACAGACATACACAGCTGTGCCAGACTCTGCTTTAAACCAAGAAACAGATGTTTAGCA 249
61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
250 AGCAGCTGCTATGAAGACTGAGAAATGTGTGGCCAAAACAAACTTGGCCAAATGGCACTTCC 309
81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
310 AGTATGATTGTGCCAAGCAACGGAACCTCTCAGCAAGCTATGAAAGGAAAGAAAGAACTG 369
101 CysValLysTyrPheGluGlnThrPheGluSerAspGlnValGluPheValGluHisLeu 120
370 TGTGTCAAAATACTTTGAGCAGTGGTCAGAGTCAGATCAGATCAAGTGAATTTGTGGAACATCTT 429
121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
430 ATATCCAAATGTGTCAATTAACCAACATGGGCACATAACTCGTATCTTAACCTATGTTG 489
141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
490 CAGAGAGATTCATAACTGCTGTCGAGCTCGGGATGGATCATATCGCTGAGAAACATT 549
161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluThrTyr 180
550 CTGTCACTGGATGCCAAATCACTATGTCTGTGAACCTTGTGTGCAAGGAATGGTAC 609
181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
610 CGAGTCACCTCTGATGGCATGCTGTGGAAGAAGCTTATCCAGAGAATGGTCAGGACAGAT 669
201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220
670 TCTCTGTGGAGAGCGCTGCGAGAACGAGAGATGGGGACAGTATTTATTCAAAACCAAA 729
221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
730 CTTCTGTGACGGGAATGCTCTCTCCAACTCTTTTATAGAGCACTTTATCCTAATAATTATA 789
241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
790 CAAGACATTGAGACAAATAGAACTTAATGGAGATGTGGAAGACATAGTTTACAGAGAATT 849
261 HisCysArgSerGlnThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
850 CACTGCGGAAAGTGAACCAAGAGAGTTTACTGTTTACAGTATGATGATCAGAAAAATA 909
281 ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300
910 GTAAGCGGCTTTCGAGACACACAAATCAGATCTCGGATATAAAACACATTTGGAATGCCAG 969

QY 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
Db 970 CGAATTTCTCAGAGCCATACAGTTTCAGTCTCTGTCTCCAGTATGATGAGAGAGTGATC 1029
QY 321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
Db 1030 ATACAGAGATCATCGGATTCACGGTCCAGAGTCAGAGTGGGATGTAATAATACAGGTGAATGCTA 1089
QY 341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
Db 1090 AACACGTTGATTACCATTTGTGAACGAGTTCTGCACCTTCCGCTTCAATTAATGGCATGATG 1149
QY 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
Db 1150 GTGACCTGCTCCAAGATCGTTCATTCGTGTATGGGATATGGCTCCCACTGACATT 1209
QY 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
Db 1210 ACCCTCCGGAGGGTGTGTCGGACACCGAGCTGCTCAATGTGTAGACTTTGATGATGAC 1269
QY 401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
Db 1270 AAGTACATTTCTTCGATCTCGGGATAGAACTATAAAGGTATGGACACCAAGTACTTGT 1329
QY 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
Db 1330 GAATTTGTAAGGACCTTAAATGGACACAAACGAGGCATTGCCCTTTTCAGTACAGGCAC 1389
QY 441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
Db 1390 AGSGTGTAGTAGTGGCTCATCTCACACACTATCAGATTATGGGACATAGAAATGTGT 1449
QY 461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
Db 1450 GCATGTTTACAGAGTGTAGAGGCCATGAGGAATTTGGTGGTGTATTTCGATTGATAAC 1509
QY 481 LysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAla 500
Db 1510 AACAGGATAGTCAGTGGGCTTATGATGAAAAATAAAGTGTGGGATCTTGTGGCTGCT 1569
QY 501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
Db 1570 TTGACCCCCGTCCTCTCGAGGACACTCTCTACAGACCCCTTGTGGAGCATTCGGGA 1629
QY 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
Db 1630 AGAGTTTTTCGACTACAGATTTGATGAATTCAGATTGTCAGTAGTTCCATCATGATGACACA 1689
QY 541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
Db 1690 ATCCTCATCTGGGACTTCTTAATGATCCAGCTGCCAAGCTGAACCCCCCGTTCCCT 1749
QY 561 SerArgThrTyrThrTyrIleSerArg 569
Db 1750 TCTCGAACATACACCTACATCTCCAGA 1776
RESULT 6
AF101784 1818 bp mRNA linear PRI 21-JAN-1999
LOCUS Homo sapiens b-TRCP variant E3RS-IkappaB mRNA, partial cds.
DEFINITION AF101784
ACCESSION AF101784.1 GI:4165135
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1818)
Yaron,A., Hatzubai,A., Davis,M., Lavon,I., Amit,S., Manning,A.M.,
Andersen,J.S., Mann,M., Mercurio,F. and Ben-Neriah,Y.
Identification of the receptor component of the
IkappaBalpha-ubiquitin ligase

AL Nature 396 (6711), 590-594 (1998)
 NE 99075339
 ED 9859996
 CE 2 (bases 1 to 1818)
 RS Yaron, A., Hatzubah, A., Mercurio, F., Mannig, A. M., Andersen, J. S.,
 Mann, M., and Ben-Neriah, Y.
 Direct Submission
 Submitted (27-OCT-1998) Immunology, Hebrew University-Hadassah
 Medical School, Jerusalem, Israel
 Location/Qualifiers
 S 1..1818
 urce /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 S <1..1818
 /note="beta-transducin repeats-containing protein;
 receptor component of IkappaBa-ubiquitin ligase; contains
 F-box and WD regions"
 /codon_start=1
 /product="b-TRCP variant E3RS-IkappaB"
 /protein_id="AAD08702.1"
 /db_xref="GI:4165136"
 /translation="MDPAEAVLQERKALFKMSPSSLWLGSSLDASMPSLRCLYNPG
 TGAATAPONSSEEDCNCEPPRKIIPEKNSURQTYNSCARILNQETVCLASTAMKT
 ENCVAKTLANGTSSMIVPKQKLSASVEKEKLCVKYFEOMSESDOEFVEHLISOM
 CHYQGHINSYKPKLQRFIATLPARGLDHIAENILSYLDAKSLCAELVCKEYRV
 TSDGMWKKLIERWRTDSLWGLAERRGQYLFKNKPPDGNAPNSFYALYFKII
 QDIETESNWRGSRISLQRIHCRSTSGVGLYDQKIVSGLRDNTIKWIKNTLE
 CKRIILHTGTSVLCIYQDERVITIGSSDSSTVRVMDVNTGEMLTILHCEAVLHLRFN
 NGMWYTCSDRSIAVMDASPTDITLRRVLVGHRAAVNVDPDDKYIVSAGSDRTIKV
 WNTSCFVRTLNGHKGRIACIQYDRDLRVVSGSSDNTIRLNDIECGACLRVLEGHEEL
 VECIRPDNKRILVSGAYDKIKWDLVLAALDPAPAGTIGLTLVHSESRVRLQDFEF
 QIVSSSHDDTIILWFLNDPAQAEPSPSIYTIYSR"
 UNT 532 a 385 c 445 g 456 t

nt Scores:
 O.: 1-25e-271 Length: 1818
 Similarity: 3006.00 Matches: 569
 cal Similarity: 94.05% Conservative: 0
 Indels: 36
 Gaps: 1

01-168B-2 (1-569) x AF101784 (1-1818)

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMet----- 16
 1 ATGGACCCGGCCGAGCGGTGTCGCAAGAGAAGGCACTCAAGTTTATGTGCTTATGCC 60
 16 ----- 16
 61 AGGTCTCTGTGGCTGGGCTGCTCCAGCCTGGCGGACAGCATGCTTCGCTGGATGCTG 120
 17 -----AsnSerSer:GluArgGluAspCys 24
 121 TATAACCCAGGACTGGCGCACTCAGAGCTTTCACGAATTCCTCAGAGAGAAAGCTGT 180
 25 AsnAsnGlyGluProProArglysllelleProGluLysAsnSerLeuArgGlnThrTyr 44
 181 AATAATGGCGCAACCCCTAGGAAGATAATACAGAGAAGAAATTCACCTTAGACAGACATAC 240
 45 AsnSerCysAlaArgLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 64
 241 AACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTAGTATGTTAGCAAGCACTGCTATG 300
 65 LysThrGluAsnGlnCysValAlaLysThrLysLeuAlaAsnGlyThrSerSerMetIleVal 84
 301 AACACTGAGAAATTGTGTGGCCAAAACAAACCTTGCCAAATGGCACTTCCAGTAGATGTG 360
 85 ProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLysValLysTyr 104
 361 CCCAAAGCAACGGAAACTCTCAGCAAGCTATGAAAGGAAAGAAAGCACTGTGTGTCAAATAC 420

QY 105 PheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeuIleSerGlnMet 124
 DB 421 TTTGAGCAGTGGTTCAGAGTCAGATCAAGTGAATTTGTGGAAACATCTTATATCCCAATG 480
 QY 125 CysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeuGlnArgAspPhe 144
 DB 481 TGTCAATTACCACATGGGCACATAAACTCGTATCTTTAAACCTATGTTGCAGAGAGATTTC 540
 QY 145 IleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIleLeuSerTyrLeu 164
 DB 541 ATTAAGTCTCTGCCAGCTCGGGATTTGGATCATATCGTGAGAACATTTCTGTACACCTG 600
 QY 165 AspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGlnTrpTyrArgValThrSer 184
 DB 601 GATGCCAAATCACTATGTGCTGCTGAACCTTGTGTCAAGGAATGGTACCGAGTGACCTCT 660
 QY 185 AspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAspSerLeuTrpArg 204
 DB 661 GATGGCATGCTGTGGAGAAGCTTATCGAGAGAAATGGTCAGGACAGATTCTCTGTGAGA 720
 QY 205 GlyLeuAlaGluArgGlyTyrGlyGlnTyrLeuPheLysAsnLysProProAspGly 224
 DB 721 GGCTGGCAGACGAAGAGGATGGGACAGTATTTATTCAAAACAAACCTCTCTGACGG 780
 QY 225 AsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIlelleGlnAspIleGlu 244
 DB 781 AATGCTCTCTCCCAACTCTTTTATAGAGCACTTTATCTCTAAAATTTATACAAGACATTGAG 840
 QY 245 ThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSer 264
 DB 841 ACAATAGAAATCTAATTTGGAGATGTGAAGACATAGTTTACAGAGAATTCCTGCCAAGT 900
 QY 265 GluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLeu 284
 DB 901 GAAACAAGCAAGAGGATTTACTGTTTACAGATGATGATCAGAAAAATAGTAACGCGCCTT 960
 QY 285 ArgAspAsnThrIleLysIleTrpAspLysAsnThrIleGluCysLysArgIleLeuThr 304
 DB 961 CGAGACAACAACAATCAAGATCTGGGATAAAAACACATTTGGAAATGCAAGCGAAATCTCACA 1020
 QY 305 GlyHisThrGlySerValLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 324
 DB 1021 GGCCATACAGTTTCAGTCTCTGCTCCAGTATGATGAGAGAGTGATCATACAGGATCA 1080
 QY 325 SerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIle 344
 DB 1081 TCGGAATTCACCGGTCAGAGTGGGATGTAATACAGGTCAAAATGCTAAACACGTTGATT 1140
 QY 345 HisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSer 364
 DB 1141 CACCATTTGAAGCAGTTCGCACCTTGCCTTTCAATAATGGCATGATGGTGACCTGCTCC 1200
 QY 365 LysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgArg 384
 DB 1201 AAAGATCGTTCATTCCTGTATGGGATATGGCTCTCCCACTGACATTAACCTCCGAGG 1260
 QY 385 ValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspLysIleValVal 404
 DB 1261 GTGCTGTGGCAGACCGAGCTGCTGTCATGTTGTAGACTTTGATGACAAGTACATTGTT 1320
 QY 405 SerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCysGluPheValArg 424
 DB 1321 TCTGCACTCTGGGATAGAACTATAAAGGTATGAAACACAAGTACTTGTGAATTTGTAAAG 1380
 QY 425 ThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAspArgLeuValVal 444
 DB 1381 ACCTTAATGGACACAAACGAGGCATTCCTGTTTCAGATACAGGACAGGCTGGTAGTG 1440
 QY 445 SerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGlyAlaCysLeuArg 464
 DB 1441 AGTGGCTCTCTGACAACACTATCAGATTATGGACATAGAAATGTGTGTGATGTTTACGA 1500

```

465 ValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsnLysArgIleVal 484
1501 GTGTTAGAGGCCATGAGGAATGGTGGCTTGATTCGATTGATAACAGAGGATATC 1560
485 SerGlyAlaTyAspGlyLysIleLysValTrpAspLeuValAlaLeuAspProArg 504
1561 AGTGGGGCTATGATGGAAAAATTAAGTGTGGGATCTGTGGCTGCTTTGACCCCGT 1620
505 AlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGlyArgValPheArg 524
1621 GTCCTCGAGGACACTCTGCTACGACCCCTTGGAGCATTCGGAAGAGTATTTTCGA 1680
525 LeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThrIleLeuIleTrp 544
1681 CTACAGTTTGATGATTCAGATTGTCAGTAGTTCACATGATGACACATCTCATCTGG 1740
545 AspPheLeuAsnAspProAlaGlnAlaGlnAlaGluProProArgSerProSerArgThrTyr 564
1741 GACTTCCATAATGATCCAGCTCCCAAGCTGAACCCCGTTCCTCCCTTCTCGAACATAC 1800
565 ThrTyrIleSerArg 569
1801 ACTCATATCTCCAGA 1815

```

```

BC027994      2308 bp      mRNA      linear      PRI 01-MAY-2002
Homo sapiens, beta-transducin repeat containing, clone MGC:40028
IMAGE:5180993, mRNA, complete cds.
BC027994
BC027994.1   GI:20380815
MGC.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2308)
Direct Submission
Submitted (08-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
Ahter,N., Ayale,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouford,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Laric,P., Legaspi,R., Maduro,Q.L.,
Masello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J.,
Pearson,R., Stantrisp,S., Thomas,P.J., Touchman,J.W., Tsurgueon,C.,
Vogt,J.L., Walker,M.A., Wecherby,K.D., Wiggins,L., Young,A.,
Zhang,L.-H. and Green,E.D.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 62 Row: e Column: 7
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4502476.
Location/Qualifiers
1..2308
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocusID:8945"

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 62 Row: e Column: 7
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4502476.
Location/Qualifiers
1..2308
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocusID:8945"

```

/db_xref="taxon:9606"
/clone="MGC:40028 IMAGE:5180993"
/tissue_type="Brain, lung, testis, adult, pooled whole"
/clone_lib="NIH MGC 115"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
38..1855
/codon_start=1
/product="beta-transducin repeat containing"
/protein_id="AAH27994.1"
/db_xref="GI:20380816"
/translation="MDPAEAVLQEKALFMCMPRSLWLGSSSLADSMPSRLCLNPG
TCALTAFOASEREDCNNGEPKRIPEKNSLRQTVNSCARCLCLNOETVCLASTAMKT
ENCVAKLIANGTSSMTVPQKLSASVEKEKELCVKVEQMSQDOVEFVHLLSQM
CHYQGHGHSYLKPMLODPITALPARGLDHIAENILSYLDAKSCAELVCKEWTRV
TSDMLWKKLIERWVRDLSLWRGLAERRGGQLFKNKPDPGNAPNSFYRLYPKII
QDIETISNWRGSRHSLOIRHCRSETSGVYCLQYDDQKIVSGLRDNTIKIWDKNTLE
CKRIITGHSVGLCIQYDERVITGSSDSTVRVMDVNTGEMLNTLHHCEAVLHFRN
NGMWTCSGSRSLAIAMVMASTPDIITLRLVLVGHRAAVNVDFDKIVISASGDRTIKV
WNTSTCFVRTLNGHKGRIACLOYRDLNVSSSDNTIRLWDIEGACLRVLEGEHEL
VRCIREDNKRIYSGAYDGKIKYVDLVAALDPRAPAGTICLRLTILVHSGRVRFLQDFE
QIVSSSHDDTILILWFLNDPAAQAEPSPRSKTYTISR"
BASE COUNT 719 a 493 c 530 g 566 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 1.73e-271 Length: 2308
Score: 3006.00 Matches: 569
Percent Similarity: 94.05% Conservative: 0
Best Local Similarity: 94.05% Mismatches: 0
Query Match: 99.08% Indels: 36
DB: Gaps: 1

US-09-601-168b-2 (1-569) x BC027994 (1-2308)
Qy 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMet----- 16
Db 38 ATGGACCGCGCGAGCGGTGCTGCAAGAGAGGCACTCAAGTTTATGTGCTCTATGCC 97
Qy 16 ----- 16
Db 98 AGGTCTCTGTGGCTGGGCTGCTCCAGCTTGGCGAGACATGCTTGGCTGGGATGCC 157
Qy 17 ----- AsnSerSerGluArgGluAspCys 24
Db 158 TATAACCCAGGACTGGCGCACTCACAGCTTTCCAGAAATCTTCAGAGAGAGAACTGT 217
Qy 25 AsnAsnGlyGluProProArgLysIleIleProGluLysAsnSerIleuArgGlnThrTyr 44
Db 218 AATAATGGCGACCCCTAGGAAGATAATACCAGAGAGAAATTCACCTTAGACAGACATAC 277
Qy 45 AsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysIleuAlaSerThrAlaMet 64
Db 278 AACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGATGATGTTAGCAAGCACTGCTATG 337
Qy 65 LysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSerSerMetIleVal 84
Db 338 AAGACTGAGAATTTGTGGCCAAACAAACCTTCCATGGCACTTCCAGTATGATTGTG 397
Qy 85 ProLysGluArgLysLeuSerAlaSerTyrGluLysGluLysGluLeuCysValLysTyr 104
Db 398 CCCAAGCAACGGAAACTCTCAGCAAGCTATGAAAGAGAAAGAAAGAACTGTGTGCAAAATAC 457
Qy 105 PheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisIleuSerGlnMet 124
Db 458 TTGACAGTGTGTCAGATCAGATCAAGTGGAAATTTGTGGAACATCTTATATCCCAATG 517
Qy 125 CysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeuGlnArgAspPhe 144
Db 518 TGTCATTACCAACATGGGCACATAACTCGTATCTTAAACCTTATGTTGAGAGAGATTTC 577
Qy 145 IleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIleLeuSerTyrLeu 164

```

578 ATRACTGCTGCGCAGCTCGGGGATGGGATCATATTGCTGAGAACATTCTGTACATCGTG 637

165 AsplalysSerLeuCyAlaAlaGluLeuValCysLysGluTrpTyrArgValThrSer 184

638 GATCCCAAAATCACTATGCTGCTGCTGAACATTGTGTGCAAGGAATGGTACCGAGTCACTCT 697

185 AspGlyMetLeuTrpLysLysLeuLeuGluArgMetValArgThrAspSerLeuTrpArg 204

698 GATGGCATGCTGTGGAGAGAGCTTATCGAGAGAAATGGTCAGGACAGATTCTCTGTGGAGA 757

205 GlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLysProProAspGly 224

758 GGCCTGGCAGAACGAGAGGATGGGACAGTATTATTTCAAAAACAAACCTCTGACGGG 817

225 AsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysLysLeuLeuAspLeuGlu 244

818 AATGCTCTCCCACTCTTTTATAGAGCACTTTATCCPAAATATATACAGACATTGAG 877

245 ThrLeuSerAsnTrpArgCysGlyArgHisSerLeuGlnArgLysHisCysArgSer 264

878 ACAATAGAAATCAATTCGAGATGTGGAAGACATAGTTTACAGAGAATTCACTGCCGAAGT 937

265 GlyThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysLeuValSerGlyLeu 284

938 GAAACAGCAAGAGAGTTTACTGTTTACAGTATGATGATCAGAAATATAGTAAAGCGCCTT 997

285 ArgAspAsnThrLysLysLeuTrpAspLysAsnThrLeuGluCysLysArgLysLeuThr 304

998 CGAGACAACAACATCAAGATCTGGGATAAAAAACACATTTGGAATGCAAGCGAATTTCTACA 1057

305 GlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValLysLeuThrGlySer 324

1058 GGCATACAGGTCAGTCCTCTGTCTCCAGTATGATGATGAGAGAGTATACACAGATCA 1117

325 SerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuLeu 344

1118 TCGGATTTCCAGGTCAGAGTGGGATGTAATACAGGTGAATGCTAAACACGTTGATT 1177

345 HisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSer 364

1178 CACCATTTGTGAAGCAGTTCTGCACATTCGCGTTTCAATTAATGGCATGATGTCACCTGCTCC 1237

365 LysAspArgSerLeuAlaValTrpAspMetAlaSerProThrAspLeuThrLeuArgArg 384

1238 AAAGATCGTTCCATGCTGTATGGGATATGCGCTCCCACTGACATTTACCTCCGGAGG 1297

385 ValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspLysTyrLeuVal 404

1298 GTGCTGTGCGACACCGAGCTGCTGTCATATGTTGTAGACTTTTATGACCAAGTACATTGTT 1357

405 SerAlaSerGlyAspArgThrLysValTrpAsnThrSerThrCysGluPheValArg 424

1358 TCTGCATCTGGGGATAGACTATTAAGGTATGGAACAACAGTACTTGTGAATTTGTAAGG 1417

425 ThrLeuAsnGlyHisLysArgGlyLeuAlaCysLeuGlnTyrArgAspArgLeuValVal 444

1418 ACCTTAATGACACAACAGGAGCTGCTGTTTTCAGTATACAGGACAGGCTGGTAGTG 1477

445 SerGlySerSerAspAsnThrLysArgLeuTrpAspLysLeuCysGlyAlaCysLeuArg 464

1478 AGTGGCTCATCTGACACACTATACAGATTATGGGACATAGAATGTGGTGCATGTTTACGA 1537

465 ValLeuGluGlyHisGluGluLeuValArgCysLeuArgPheAspAsnLysArgLysLeuVal 484

1538 GTGTTAGAGCCCATGAGGAATGGTGGTGTGTTATTCGATTGATTAACAGAGGATAGTCT 1597

485 SerGlyAlaTyrAspGlyLysLysLeuValTrpAspLeuValAlaAlaLeuAspProArg 504

1598 AGTGGGCGCTATGATGAAAAAATTAAGTGTGGGATCTTGTGGCTGCTTTGGACCCCGCT 1657

505 AlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGlyArgValPheArg 524

1658 GCTCTTTCAGGGACACTCTGTCTACGGACCCCTTGTGGAGCATTTCCGGAAGAGTTTTCGA 1717

Qy 525 LeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThrIleLeuIleTrp 544

Db 1718 CTACAGTTTGATGAATTCAGAGATTGTAGTAGTTTCACATGATGACACAAATCTCTCATCTGG 1777

Qy 545 AspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerProSerArgThrTyr 564

Db 1778 GACTTCTTAATGATCAGCTGCCCAAGCTGAACCCCGTTCCCTCTCTCGAACATAC 1837

Qy 565 ThrTyrIleSerArg 569

Db 1838 ACCTACATCTCCAGA 1852

RESULT 8

LOCUS AF099932 1712 bp mRNA linear ROD 21-JAN-1999

DEFINITION Mus musculus beta-TrCP protein E3RS-IkappaB mRNA, complete cds.

ACCESSION AF099932

VERSION AF099932.1 GI:4008019

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1712)

AUTHORS Yaron,A., Hatzubai,A., Davis,M., Lavon,I., Amit,S., Manning,A.M., Andersen,J.S., Mann,M., Mercurio,F. and Ben-Neriah,Y.

TITLE Identification of the receptor component of the IkappaBalpha-ubiquitin ligase

JOURNAL Nature 396 (6711), 590-594 (1998)

MEDLINE 98075339

PUBMED 9859996

REFERENCE 2 (bases 1 to 1712)

AUTHORS Yaron,A., Hatzubai,A., Mercurio,F., Manning,A.M., Andersen,J.S., Mann,M. and Ben-Neriah,Y.

TITLE Direct Submission

JOURNAL Submitted (19-OCT-1998) Immunology, Hebrew University of Jerusalem, Ein Karem, Jerusalem 91120, Israel

FEATURES

source

1. .1712

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

1. .1710

/notes="beta-transducin repeat-containing protein, receptor component of IkBa-ubiquitin ligase"

/codon_start=1

/product="beta-TrCP protein E3RS-IkappaB"

/protein_id="AAD08701.1"

/db_xref="GI:4008020"

/translation="MDPAEAVLOEKALFMNRSBEDCNGEPPPKIIPKNSLRQTY NSCARLINCQETVCLTSTAMKTENCYAKAKLANGTSSMIVPKQKLSASYEKELCV KTFQWESDQVEFVHELISQMHCHYQKHLINSLKPMLOQDFITALPARGDQHTAENI LSYLDASQAEALVCKEWMRYVTSQMLNKLIERMVRTSLMRGLAERRGQGLVFK NKPPDNAPNSFYALYPKTIQDIETTESNRCGRHSLOTHCRSETSKGYCLOYDV DQKIVSGLDNTIKWDSLTCKELTGTGSLVCLQYDERVITITSSDSTVRVWDV NAGEMLTUHVHCEAVLHLPFNNMVTCSKDRSAVMDASPTDITLRRVLVGHRAA VNVVDKDKIVYASGSDRTIKVNTSTCFVRLTNGHKGAGIACLOYRDLRVVSGSDN TIRLMDIEGACLRVLEGEHELVRIRDNKLIVSGAYDGKIKWMDLMAALDPRAPAG TICRLTIVSHSRVFRLOQDFEQIVSSSHDDTLILWDFLNDPFAHAEPSPSRITYTISR"

BASE COUNT 469 a 399 c 453 g 391 t

ORIGIN

Alignment Scores: 8e-271 Length: 1712

Fred. No.: 2997.00 Matches: 561

Score: 99.12% Conservative: 3

Percent Similarity: 98.59% Mismatches: 5

Best Local Similarity: 98.78% Indels: 0

Query Match: 10 Gaps: 0

DB:

US-09-601-168B-2 (1-569) x AF099932 (1-1712)


```

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
1 ATGGACCGCGCGCGCGGTGTCGAGGAGAAGCGCTTAAGTTTATGAATTCCTCAGAG 60
21 ArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLysAsnSerLeu 40
61 AGAGAAGACTGTAATAATGGCGAACCCCTAGGAAGATATATACGAGAGAAGATTCATT 120
41 ArgGlnThrThrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
121 AGACAGACTTACAACAGCTGTGCCAGGCTTTGCATAAACCAGAGACAGATATGCTAACA 180
61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
181 AGCACTGCTATGAAGACTGAAATTTGTGGCCAAAGCCAACTTGCCTAATGGCACTTCC 240
81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
241 AGCATGATTGTGCCCAAGCAGCGGAACTCTCAGCAAGCTATGAGAAGAAAGGAGCTG 300
101 CysValLysTyrPheGluGlnThrSerGluSerAspGlnValGluPheValGluHisLeu 120
301 TGTGTCAGTATTTCAGCAGTGGTCAGAGTCTGATCAAGTGGAAATTTGTAGAACACCTT 360
121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
361 ATATCCCAATGTGTCACTACCAAGCTGGGACATCAACTCTTACTTAACCTATAGCTG 420
141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
421 CAGAGGATTTTCATNACTGCACTGCCAGCAGCAGGGGTCTGGACCATCGCTGAGAACAT 480
161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluThrPyr 180
481 CTGTCAATCTTGGACGCCAAGTCACTGTGTGTCTGTGAGCTCGTGTCAAGGAATGGTAC 540
181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
541 CGCGTCAGCTCGGACGGCATGCTGTGGAAAAAGCTCATCGAGAGGATGGTCAGGACGAC 600
201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220
601 TCTCTGTGGCGAGCGCTGGCAGAGCGCAGAGGCTGGGGACAGTACTTATTAAAAACAAA 660
221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
661 CTCTCTGATGAGAACGCTCTCCCACTCTCTTTATAGACGCTTTATCCTTAATCATTA 720
241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
721 CAAGACATTGAGACAATAGAGTCCAATGGAGATGTGGCGACATAGATTACAGAGAATC 780
261 HisCysArgSerGlnThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
781 CACTGCGGAGTGAACAAGTAAAGGGGTTTACTGTTTACAGTACGACACAGAGATA 840
281 ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300
841 GTCAGCGGCTTCGAGACACACATCAAGATCTGGGATAAAGACACATCGAATGCCAAG 900
301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
901 CGGATTCTCAGCGGCCACAGCGGCTCCGCTGCTGTCGAGTACGATGAGAGGGGTGATC 960
321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
961 ATCAGAGGCTCCCTCAGACTCCACCGTCAGAGTGTGGATGTAATCCAGGTGAGATGCTA 1020
341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
1021 AACACATTGATTACCACTGTGAAGCGGCTTCGCACCTGGGCTTCAATATGGCATGATG 1080

```

361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
1081 GTGACCTGTTCAAAGACCGTTCCATCGCTGTGTGGGATATGGCTTCCCACACTGACATC 1140
381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
1141 ACCCTCAGGAGGGTGTGTGTGGGACACCGAGCTGCGGTCAATGTTGTAGACTTTGATGAC 1200
401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
1201 AAGTACATCGTTTCTGCTCTGAGATAGAACATAAAGGTGTGGAACACAAGTACCTGT 1260
421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
1261 GAATTGCTAAGACCTTAANTGGGCACAAAGCTGGCATCGCTGTTTGCAGTACAGAGAC 1320
441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
1321 AGGCTGCTGTGAGCGGCTCTCTGCACAACACCATCAGCTGTGGGACATAGAGTGTGGA 1380
461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
1381 GCATGCTTCGAGTGTGTGGAGGGCCATCAGGAGTGGTACGCTGCATTCGATTTCATTAAC 1440
481 LysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAla 500
1441 AAAAGATAGTGAGCGGAGCCCTATGATGGGAAATTAAGTGTGGATCTTATGCTGCT 1500
501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
1501 TTGACCCCGCTGCTCCAGCAGGACTCTCTGCTGTGGGACACTTGTGGAGCAATCTCGA 1560
521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerSerHisAspAspThr 540
1561 AGATTTTTCGCTCCAGTTTGTGAATTCAGATTTGCAGTAGTTCATATGATGACACA 1620
541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
1621 ATTCTCATCTGGGACTTCTGTAATGATCCAGCTGCTCAGCTGAACCGCCGCTCCCT 1680
561 SerArgThrTyrThrTyrIleSerArg 569
1681 TCTCGACATACACCTTACATCTCCAGA 1707

BC003989 2081 bp mRNA linear ROD 16-APR-2003
Mus musculus beta-transducin repeat containing protein, mRNA (cDNA
Clone MGC:7517 IMAGE:3491843), complete cds.
BC003989
BC003989.1 GI:13278339
MGC.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2081)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Sapich, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,
Schetz, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.O., Malek, J.A., Gunaratne, P.H., Richards, S.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Heltton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Maira, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22398257
AL NE
ED CE
2 (bases 1 to 2081)
Direct Submission
Strausberg, R.
Submitted (28-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
K NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 8 Row: 0 Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6753209.
Location/Qualifiers

S

```

1..2081
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="MGC:7517 IMAGE:3491843"
/tissue_type="Mammary tumor metastasized to lung. Tumor
arose spontaneously from a senescent normal mammary
(clone) outgrowth infected with the virus MMV."
/clone_lib="NCI CGAP_Lu29"
/lab_host="DH10B"
/notes="Vector: PCMV-SPORT6"
1..2081
/gene="Btrc"
/notes="synonyms: FWD1, Slimb, b-TrCP, mSlimb"
/db_xref="LocusID:12234"
/db_xref="GI:13278340"
7..1716
/codon_start=1
/product="beta-transducin repeat containing protein"
/protein_id="AAH03989.1"
/db_xref="GI:13278340"
/db_xref="LocusID:12234"
/translation="MDPAEVLQEKALPMNSSEDCNNGEPKRIPEKNSLRQV
NSCARLCINQETVCLTSTAMKTCVAKLANCTNSIVPKQKLASVEKEKELCV
KFEQKSDQVEFVHLISQCHTQGHILSYLPMQLQRFITAPALPARGLDHIAENI
LSYLAQSKCAELVCKEWRYVSDGMLWKKLIETRWETSLMRGLAERGWGYLFE
NKPPENAPPNSFYALPYKIIQDIETIESNWRGHSLOIHCRTSKGVCLQYD
DQKIVSGLRDNTIKWXSLECKEILHTGTSVLQYDERVITGSSDSVTRWDV
NAGEMLNTLIHCEAVLHLRFNNGMWTCSKORSIAVWDMASPTDITLRRVLVGRRA
VNVDPDDKVIIVSAGSDRTIKVWNTSTCFVTRLNKHGKGIACIQYDRDLVSSDN
TIRLWDGCAQLVLEHGEELVRCIRFNKRVSGADYDGKIKVMDLMAALDPPAPAG
TICLRDLVHSGSRFRLQDFEQIVSSSHDDTILWDFLNDPAHAHSPSPSRITYT
ISR"
JNT 568 a 502 c 529 g 482 t

```

at Scores:

1.05e-270 Length: 2081
2997.00 Matches: 561

```

Percent Similarity: 99.12% Conservative: 3
Best Local Similarity: 98.59% Mismatches: 5
Query Match: 98.78% Indels: 0
DB: 10 Gaps: 0
US-09-601-168B-2 (1-569) x BC003989 (1-2081)

Qy 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
Db 7 ATGACCCGCGCAGAGCGGTGCTGCAGGAGAAAGCGCTTAAGTTTATGATTCCTCAGAG 66
Qy 21 ArgGluAspCysAsnAsnGlyGluProArgLysIleIleProGluLysAsnSerLeu 40
Db 67 AGAAGAACTGTTAATATGCGAAACCCCTAGGAGATTAATACACAGAGAAAGATTCAC 126
Qy 41 ArgGluThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
Db 127 AGACAGACTTACACAGCTGTGCCAGGCTTGCATTAACCAAGACAGACAGTATGCTPA 186
Qy 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
Db 187 AGCACTGCTATGAAGACTGAAAATTTGTGTGGCCAAAGCCAAACTTGCCAATGSCACT 246
Qy 81 SerMetIleValProLysGlnArgLysIleSerAlaSerTyrGluLysGluLysGluLeu 100
Db 247 AGCATGATGTGCTCCCAAGCAGCGGAACTCTCAGCAAGCTATGAGAGGAAAGAGAGTG 306
Qy 101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
Db 307 TGTGTCAGGTATTTTTCAGCAGTGTGCAGAGTCTGATCAAGTGAATTTGTAGAACCTT 366
Qy 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
Db 367 ATATCCCAATGTGCTACCTACGAGCATGGGCACTCACTCTCTATAAACTATGCTG 426
Qy 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
Db 427 CAGAGGATTTTCACTAAGTGCATGCCAGCAGCGGGTCTGGACCAACATCGCTGAGA 486
Qy 161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 180
Db 487 CTGTCATCTTGACGCCCAAGTCACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 546
Qy 181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
Db 547 CGCGTACGTCGACCGCATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 606
Qy 201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220
Db 607 TCTCTGTGGCGAGGCTTGGCAGAGCGCAGAGGCTGGGGACAGTACTTATCAAAAACA 666
Qy 221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
Db 667 CCTCTCTGATGAGAACGCTCTCTCCCACTCTTTTATAGAGCGCTTTATCTATAAATCA 726
Qy 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
Db 727 CAAGACATTGAGACAATAGAGTCCCAATGAGATGTGGGCGACATAGTTTACAGAAATC 786
Qy 261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
Db 787 CACTGCCGAGGTGAAACAAAGTAAGGGGTTTACTGTTTACAGTACGACGACGAGAGATA 846
Qy 281 ValSerGlyLeuArgAspAsnThrIleLysIleTyrAspLysAsnThrLeuGluCysLys 300
Db 847 GTCAGGCGGCTTCGAGACAACACCATCAAGATCTGGGATAAAGACACACTGGAATGCA 906
Qy 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
Db 907 CGGATTCTCAGGGGCCACACGGGCTCCGTCCTGTGTGTGTGTGTGTGTGTGTGTGTGT 966
Qy 321 IleThrGlySerSerAspSerThrValArgValTyrAspValAsnThrGlyGluMetLeu 340

```

967	ATCAGAGGCTCCTCAGACTCCACCGTCAGAGTGGGATGTAATGCGAGGTGAGATGCTA	1026	
341	AsnThrLeuIleHisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet	360	
1027	AACACATTGATTCCACCACTGTGAGCGGCTTCGACCTCGCTTCAATAATGGCATGATG	1086	
361	ValThrCysSerIleAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle	380	
1087	GTGACCTGCTCCAAAGACCGTTCCATCGCTGTGTGGGATATGGCTTCCCAACTGACATC	1146	
381	ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp	400	
1147	ACCCTCAGAGGGTCTGTGTGGGACCGGCTGGGTCGGAATGTTGAGATTTGATGAC	1206	
401	LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys	420	
1207	AAGTACATCGTTTCTGCCTCTGGAGATAGAACCAATAAGGTGTGGACACAGTACCTGT	1266	
421	GluPheValArgThrLeuAsnGlyHisIleArgGlyIleAlaCysLeuGlnTyrArgAsp	440	
1267	GAATTCTGAAGACCTTAATGGGCACCAAGCGTGGCATCGGCTGTTTGCAGTACAGAC	1326	
441	ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAsnIleGluCysGly	460	
1327	AGGCTGTGTGGCGGCTCTCTGACAAACCATCAGGCTGTGGACATGAGTGTGGA	1386	
461	AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn	480	
1387	GCATGCTCGGAGTGTGGAGGGCCATCAGGAGTGTGTACGCTGCATTCGATTGATAAC	1446	
481	LysArgIleValSerGlyValaTyrAspGlyLysIleLysValTrpAspLeuValAlaAla	500	
1447	AAAAGATAGTAGGCGGAGCCTATGATCGGAAATTAAGTGTGGGATCTTATGCGTCT	1506	
501	LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly	520	
1507	TGGACCTCGTCTCCAGCAGGACTCTCTGTCTGGGACACTGTGGAGCATCTGGA	1566	
521	ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr	540	
1567	AGAGTTCCTCCGCTCCAGTTTATGATTCAGATTCAGATTCAGATTCAGATTCAGAC	1626	
541	IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro	560	
1627	ATTCATCTGGAGCTCTCTGAATGATCCAGCTGCTCAGCTGAACCGCCGCTCCCT	1686	
561	SerArgThrTyrThrTyrIleSerArg	569	
1687	TCTGGACATACACTACATCTCCAGA	1713	
0			
AF081887	2175 bp	mRNA	linear
ON	Mus musculus ubiquitin ligase FWD1	mRNA	complete cds.
DN	AF081887		
AF081887.1	GI:4336326		
3	Mus musculus (house mouse)		
SM	Mus musculus		
IE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
IS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
1	(bases 1 to 2175)		
IS	Hatakeyama, S., Kitagawa, M., Nakayama, K., Shirane, M., Matsumoto, M.,		
IS	Hattori, K., Higashi, H., Nakano, H., Okumura, K., Onoe, K., Good, R.A.		
IS	and Nakayama, K.-I.		
IS	Ubiquitin-dependent degradation of IkappaBalpha is mediated by a		
IS	ubiquitin ligase Skp1/Cul 1/F-box protein FWD1		
IS	Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3859-3863 (1999)		
IE	99199275		
IE	10097128		
IE	2 (bases 1 to 2175)		
IE	Hatakeyama, S. and Nakayama, K.-I.		
S	Direct Submission		
Submitted (04-AUG-1998) Department of Molecular and Cellular			
Biology, Medical Institute of Bio-regulation, Kyushu University,			
3-1-1, Maedashi, Higashi-Ku, Fukuoka 812-8582, Japan			
Location/Qualifiers			
1. 2175			
/organism="Mus musculus"			
/mol_type="mRNA"			
/db_xref="taxon:10090"			
91..1800			
/note="similar to beta-transducin repeat-containing			
protein Trp; F-box protein; contains WD40 repeats"			
/codon_start=1			
/product="ubiquitin ligase FWD1"			
/protein_id="AAD17755.1"			
/db_xref="GI:4336327"			
/translation="MDPAEVLQEKALFMSSREDGNGEPPPKIIPKNSLRQTY			
NSCARLCSNQETVCLTSTAMKTEHCVAKLANGYSSMIVPKORKLSAYEKELCV			
KIPFQSSDQVEFVEHLISQHGHIHSLQVLMQVDFITALPARGLDHIAENI			
LSYDAKSLCAABLVCBMYRVTSQMLKLIERMVTDLSLWRGLAERGNGYLFK			
NKPPDENAPPSFYRALYPKITQDIETIESNWRGHSQRIHCRSETSKGYCLQYD			
DOKIVSLGRDNTIKWDKSTLEKRLITGHTGSLQVDERVITGSSDSSTVRVWDV			
NAGEMLTLIHCEAVLHLRNNMMVTCSKDRSTAVWDMASPTDITLRLVLGHRAA			
VNVVDFDKYIVSASGDRITKVNWTSTCEFTVTLNGHKGRIACIQRYRLVVSQSDN			
TIRLMDICGACIRVLEGEELVRIRFDNKRIVSGAYDGKIKWMDLMAALDPRAPAG			
TICLRLTLEHSGRVFRLOQDFEIVSSSHDDTILIDFLNDFANAEPPSPSRITY			
ISR"			
BASE COUNT	589 a	528 c	567 g 491 t
ORIGIN			
Alignment Scores:			
Pred. No.:	1,116-270	Length:	2175
Score:	2997.00	Matches:	561
Percent Similarity:	99.12%	Conservative:	3
Best Local Similarity:	98.59%	Mismatches:	5
Query Match:	98.78%	Indels:	0
DB:	10	Gaps:	0
US-09-601-168B-2 (1-569) x AF081887 (1-2175)			
QY	1	MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu	20
Db	91	ATGGACCCGCGAGCGGCTGTGAGGAGAAAGCGCTTAAGTTTATGAAATTCCTCAGAG	150
QY	21	ArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLysAsnSerLeu	40
Db	151	AGAGAAGACTGTAATAATGGCAACCCCTAGGAAGATAATACCAAGAGAATTCACCT	210
QY	41	ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla	60
Db	211	AGACAGACTTACACAGCTGTGCCAGGCTTCATATAACCAAGACAGACAGTATGTCTACA	270
QY	61	SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer	80
Db	271	AGCAGTGTATGAGACTGAAATTTGTGTGGCCAAAGCCAACTTGCCAAATGGCACTCC	330
QY	81	SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu	100
Db	331	AGCATGATTGTGCCAAGCAGCGGAACTCTCAGCAAGCTATGAGAAGAAAGGAGCTG	390
QY	101	CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu	120
Db	391	TGTGTCAAGTATTTTGTAGCAGTGTGTCAGAGTCTGATCAAGTGGAAATTTGTAGAAC	450
QY	121	IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu	140
Db	451	ATATCCCAATGTGTCTACCTACGAGTGGGACATCACTCTACTATAAACCTATGCTG	510
QY	141	GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle	160
Db	511	CAGAGGATTTTCATAACTGCATGCGCAGCAGCGGGTCTGGACACATCGCTGAGAAT	570
QY	161	LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr	180

571 CTGTCACTACTGGACGCCAAGTCACCTGTTGTGCTGCTGAGCTTGTTGCACAGAAATGTTAC 630

181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200

631 CGCGTGAAGTCGACGCGCATGCTGTGGAAAAAGCTCATCAGAGAGTGGTCAGGACGGAC 690

201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyVTrpGlyGlnTyrrLeuPheLysAsnLys 220

691 TCTCTGTGGCGAGCCCTGGCAGAGCCGACAGCGCTGGGACAGTACTTATTCAAATAACAAA 750

221 ProProAspGlyAsnAlaProProAsnSerPheTyrrArgAlaLeuTyrrProLysIleile 240

751 CCTCTGATGAGAAGCTCTCTCCCAACTCTCTTTATAGAGCGCTTTATCTTAATCATATA 810

241 GluAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGluArgIle 260

811 CAAGACANTGAGACAATAGAGTCCCAATTGGAGATGTGGCGCACATAGTTTACAGAGAATC 870

261 HisCysArgSerGluThrSerLysGlyValTyrrCysLeuGlnTyrrAspAspGlnLysIle 280

871 CACTGCCGGAGTGAAACAAGTAAGGGGTTTTACTGTTTACAGTAGCAGCACCAGAGATA 930

281 ValSerGlyLeuArgAspAsnThrIleLysIleIleTrpAspLysAsnThrLeuGluCysLys 300

931 GTACAGCGCCCTTCGAGACAACACATCAAGACTCTGGGATAAAAGCACACTGGAAATGCAAG 990

301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrrAspGluActrValIle 320

991 CGGAATCTCAGCGGCCACACGGGCTCGGCTCTGTGTCTGACGTACGATGAGAGGGGTGATC 1050

321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340

1051 ATCACAGGCTCTCTCAGACTCCACCGCTCAGAGTGTGGGATGTAAATGCAGGTGAGATGCTA 1110

341 AsnThrIleulleHisHisCysGluAlavalLeuHisLeuArgPheAsnAsnGlyWetMet 360

1111 AACCATTTGATTCACCACTGTGAAGCGGTTGTGCACCTCGCTTCCAATAATGGCATGATG 1170

361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380

1171 GTGACCTGTTCCAAAGACCGTTCCATCGCTGTGTGGGATATGGCTTCCCCAACCTGACATC 1230

381 ThrLeuArgArgValLeuValGlyHisArgAlaAlavalAsnValValAspPheAspAsp 400

1231 ACCCTCAGAGAGGGTGTGGTGGGACACCGAGCTCGGCTCAATGTGTAGACTTTTGATGAC 1290

401 LysTyrrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420

1291 AAGTACATCGTTTCTGGCTCTGGAGATAGNACCATTAAGGTGTGGAAACACAGATGACCTGT 1350

421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrrArgAsp 440

1351 GAAITCGTAAGGACCCCTAAATATGGGCAACAGCGTGGCATCGCTGTTTTCAGTACAGAGAC 1410

441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460

1411 AGCGTGTGTGACGGCTCTCTGTGACACACCATCAGGCTGTGGGACATGAGTGTGGA 1470

461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480

1471 GCATCGCTCGCAGTGTGGAGGGCCATGAGGAGTGTGATCGCTGATTCGATTTGATAAC 1530

481 LysArgIleValSerGlyAlaTyrrAspGlyLysIleLysValTrpAspLeuValAlaAla 500

1531 AAAAGGATGATGACGGAGGCTATGATGGGAAAATTTAAAGTGTGGGATCTTTATGGCTGT 1590

501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520

1591 TTGGACCGCGTGTCTCCAGCAGGAGACTCTCTGTCTGCGGACACTTGTGGAGATCTTGGA 1650

521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540

1651 AGAGTTTTCCGCCCTCCAGTTTGAATTCAGATTCCAGATTGTCAATGTTTCAATGATGACACA 1710

Qy 541 IlleJeu1lTrpaspPheLeuAenAspProAlaAlaGlnAlaGlnProProAArgSerPro 560
 Db 1711 ATTCTCATCTGGGACATCTCTGGAATGATCATCGCTGCTCAGCTGAACGGCCGCTCCCT 1770
 Qy 561 SerArgThrTyrThrTyrIleSerArg 569
 Db 1771 TTCGGACATACACCTACATCTCCAGA 1797

RESULT 11
 AF112979 1710 bp mRNA linear ROD 02-MAR-1999
 LOCUS Mus musculus beta-transducin repeat containing protein mRNA,
 DEFINITION complete cds.
 ACCESSION AF112979
 VERSION AF112979.1 GI:4140717
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Mus.
 REFERENCE 1 (bases 1 to 1710)
 AUTHORS Spencer R., Jiang J. and Chen, Z.J.
 TITLE Signal-induced ubiquitination of IkappaBalpha by the F-box protein
 SImb1-induced ubiquitination of IkappaBalpha by the F-box protein
 JOURNAL Genes Dev. 13 (3), 284-294 (1999)
 MEDLINE 99145465
 PUBMED 9990853
 REFERENCE 2 (bases 1 to 1710)
 AUTHORS Chen, Z.J.
 TITLE Direct Submission
 JOURNAL Submitted (10-DEC-1998) Molecular Biology and Oncology, UT
 Southwestern Medical Ceter, 5323 Harry Hines Blvd, Dallas, TX
 75235-9148, USA

FEATURES
 source Location/Qualifiers
 1..1710
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 1..1710
 /function="ubiquitinates phosphorylated IkBa"
 /notes="beta-TrCP; F-box protein; IkB-ubiquitin ligase;
 substrate recognition subunit of SCF complex; similar
 Homo sapiens beta-TrCP and Drosophila melanogaster Slimb"
 /codon_start=1
 /product="beta-transducin repeat containing protein"
 /protein_id="AA004181.1"
 /db_xref="GI:4140718"
 /translation="MDPAEAVLQEKALKTFMNSEREDCNNGPPRKILPEKNSLRQTY
 NSCARLINCQETVLTSTAMKTENCAKLANGTSMIVPORKLSASYEKEKLCV
 KYFEQSDQVEFVHLLSQMCHYQGHINSYLKEPQLQDFITALPARGDLIAENI
 LSVLDKSLCAAEVLCKVWRVTSQGLWKKLIERVMVTDLSLWRGLAERRGMQGYLFP
 KPPDENAPNSFYRALPKILODITIESNWRCGRHSIQRIHCRSETSGVYCVLOY
 NQVIGSLDNTIKINDKSTLECKRLTCTGTSVLCGLGERVILTGSDSDTVVMQV
 NAGENMLTIHHCCEVHLRFNNGMMVTGSKORSIAVWNASPTDIILRLVLVGHRA
 VNVVDDDKIVSASGDRITKWNSTCFEVRTLVGKRGIACTQDFRLVVGSSDGN
 TIRLWDIEGACLRVLEGEELVRCIFDNKTRIVSGAYDGKIKVMDMAALDPRAPAG
 TGLRLTLVHSGRVFLQDFEFOIVSSHSDDTILIWDFLNDPAAHAEPPSPSRRTTY
 ISR"

BASE COUNT 469 a 399 c 453 g 389 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3,62e-270 length: 1710
 Score: 2990.00 Matches: 560
 Percent Similarity: 98.95% Conservative: 3
 Best local Similarity: 98.42% Mismatches: 6
 Query Match: 98.55% Indels: 0
 DB: 10 Gaps: 0

US-09-601-168B-2 (1-569) x AF112979 (1-1710)
 1 MetAspProAlaGlnAlaValLeuGlnGluValAlaLeuIvsPheMetAenSerSerGlu 20
 Ov

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20

```
1 ATGGACCCGGCAGAGCGGTGTCGAGGAGAAAGCGCTTAAGTTTATGAATTCCTCAGAG 60
21 ArgGluAspCysAsnAsnGlyClnProArgLysLleIleProGluLysAsnSerLeu 40
61 AGAGAAGACTGTAAATAATGGCGAACCCCTAGGAAGATAATACCAGAGAAATTCACCT 120
41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
121 AGACAGACTTACACAGCTGTGCAGGCTTTGCATAAACCAAGAGACAGATATGCTAACA 180
61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
181 AGCACTGTATCAGACTCAAAATTTGTGGCCAAAGCCAAATGTCCTGCAATGGCACTTCC 240
81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
241 AGCATGATTTGTGCCCAAGCAGCGGAACTCTCAGCAAGCTATGAGAGGAAAGAGAGCTG 300
101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
301 TGTGTCAAGTATTTTGGCAGTGTGTCAGAGTCTGATCAAGTGGAAATTTGTAGAACACCTT 360
121 IleSerGlnMetCysHisTyrClnHisGlyHisLleAsnSerTyrLeuLysPheMetLeu 140
361 ATATCCCAAAATGTCTACTACAGCATGGGCACATCAACTCTCTACCTAAAAACCTATGCTG 420
141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisLleAlaGluAsnIle 160
421 CAGAGGGATTTCTAACTGCACTGCCACAGCGGGTCTGGACCACTGCTGAGAACTT 480
161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 180
481 CTGTCTACTTTGGACGCCCAAGTCACTGTGTGCTGCTGAGCTGTCGCAAGAAATGGTAC 540
181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
541 CGCTGTGACGTGGACGGGATGTGTGGAAAAAGCTCATCGAGAGGATGCTCAGGACGGAC 600
201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220
601 TCTCTGTGGCGAGGCTGGCAGAGCGCCAGAGCTGGGACAACTACTTATCAAAAACAAA 660
221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
661 CCTCTGTATGAGAACCGCTCTCCCAACTCTTTATAGAGCGCTTATCTCTAAAATCATA 720
241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
721 CAAGACATTGAGACAAATAGAGTCCAAATTTGGAGATGTGGCGACATAGTTTACAGAGAATC 780
261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
781 CACTCCGGAGTGAACCAAGTAAAGGGTTTACTGTTTACAGTACGACGACACAGAGATA 840
281 ValSerGlyLeuArgAspAsnThrIleLysLleTrpAspLysAsnThrLeuGluCysLys 300
841 GTCAAGCGGCTTCGAGACAAACACCATCAAGATCTGGGATAAAAGCACACTGGAATGCAAG 900
301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
901 CGGATTTCTACGGGCCACACGGGCTCCGTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
961 ATCAGAGGCTCTCAGACTCCACCTCAGAGTGTGGGATTAATTCAGAGGTGAGATGCTA 1020
341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
1021 AACACATTGATTCACCACTGTGAAGCGGTTCTGCACCTGCGCTTCAATAATGGCATGATG 1080
361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
```

```
1081 GTGACCTGTTCCAAAGACCGTTCCATCGCTGTGTGGGATATGGCTTCCCAACTGCACATC 1140
381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
1141 ACCCTCAGAGGGTGTGTGTGGGACACCGAGTGGGTCAATGTTGTAGACTTTTGTATGAC 1200
401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
1201 AAGTACATCGTTTCTGCCTCTGGAGATAGAACCAATAAAGGTGTGGACACACAAGTACCTGT 1260
421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
1261 GAATTCGTAAAGACCCCTAAATGGGCACAAAGCGTGGCATCGCTGTTTGCAGTACAGAGAC 1320
441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
1321 AGCTGTGTGTGAGCGCTCTCTGACAAACACCATCAGGCTGTGGACATAGAGTGTGA 1380
461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
1381 GCATGCTCGAGTGTGGAGGCGCATGAGGAGTTGGTACGCTGCATTCGATTTCGATAAC 1440
481 LysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuValAlaIle 500
1441 AAAAGGATAGTGGCGGAGCCCTATGATGGGAAAAATTTAAAGTGTGGATCTTATGGCTGCT 1500
501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
1501 TTGGACCCGCTGCTCCAGCAGGACTCTCTGTCTGGGACACTTGTGGAGCATTCCTGA 1560
521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerSerHisAspThr 540
1561 AGAGTTTCCGCTCCAGTTTGTGATGAATTCAGATTGTCAGTAGTTCACATGATGACACA 1620
541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
1621 ATTCTCATCTGGACTTCTGATGATCAGCTGTCTCAGCTGTAACCGCCGCTCCCT 1680
561 SerArgThrTyrThrTyrIleSerArg 569
1681 TCTCGACATACCTACATCTCCAGA 1707

RESULT 12
AF110396 1979 bp mRNA linear ROD 28-JUN-1999
LOCUS Mus musculus beta-transducin repeat-containing protein mRNA,
DEFINITION complete cds.
ACCESSION AF110396
VERSION AF110396.1 GI:5230821
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1979)
AUTHORS Winston,J., Ellledge,S.J. and Harper,J.W.
TITLE Direct Submission
JOURNAL Submitted (01-DEC-1998) Biochemistry, Baylor College of Medicine,
One Baylor Plaza, Houston, TX 77030, USA
FEATURES
source 1..1979
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
CDS 1..1710
/notes="beta TRCP"
/codon_start=1
/product="beta-transducin repeat-containing protein"
/protein_id="AA041025.1"
/db_xref="GI:5230822"
/translation="MDPAEAVLQEKALPKMNSRERDCNNGEPKPIIPEKNSLRQTY
NSCARLINCQETVCLTSTAMKLVNCAKALANGTSSMIVPKRKLASYSKEKELCV
KYFQESDSDQVEFVHLSIQMCHYQGHINSYLPKMLQRFITALPARGLDHITENI
```

LSYLDKSLCAAELVCKEWRVTSQMLWKLLIERMVRITSLMRGLAERRGWOYLFK
 NKPPDENAPPNSFYALXPKIOTDITTESNRCGRHSIORHCRSETSGKVCLOYD
 DOKIVSGERDNTIKTWKSTLECKELTGHGTGSLVLCLOYDERVLIITGSSDSTVRVMDV
 NAGEMDLTILTHCEAVLHLEFRNNMNYTCKSDRSIAVWDMASPTDILRLRVLGHRAA
 VNVDFDDKTIIVSAGSDRTIKVWNTSTCEFKVRLTNGHGLACLOYDRIVVSSGSDN
 TIRLWDIECGACLRVLEGEHELVRICFRDNKRIIVSGRYDGKIKYWDMLAALDPRAPAG
 TILCLRTLVEHSGRVERFLQDFEQIVSSSHDDTILITWDFLNDPAHAHPPRSPSRITY
 ISR"

UNT 535 a 480 c 516 g 448 t

nt Scores:

O.: 5.88e-269 Length: 1979
 2978.00 Matches: 557
 Similarity: 98.59% Conservative: 4
 cal Similarity: 97.89% Mismatches: 8
 atch: 98.15% Indels: 0
 10 Gaps: 0

31-168B-2 (1-569) x AF110396 (1-1979)

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
 1 ATGGACCCGGCAGAGCGGTGCTGACAGAGAAAGCGCTTAAGTTATGAATTCCTCAGAG 60
 21 ArgGluAspCysAsnAsnGlyGluProArgLysLlelleProGluLysAsnSerLeu 40
 61 AGAGAAGACTGTAATAATGCGAACCCCTAGAGAGATAATACACAGAGAAGAAATTCACCT 120
 41 ArgGlnThrTyAsnSerCysAlaArgLeuLysLeuAsnGlnGluThrValCysLeuAla 60
 121 AGACAGACTTACACAGCTGTGCCAGCTTTTCATTAACCAAGACAGAGTAGTGTCTAACA 180
 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
 181 AGCACTGCATGAAGACTGAAATTTGTGTGCCAAGACCAAACTTGCCAAATGGCACTTCC 240
 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyArgGluLysGluLeu 100
 241 AGCATGATTGTGCCAAGCAGCGGAACCTCTCAGCAAGCTATGAGAAGAAAGAGAGCTG 300
 101 CysValLysTyPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
 301 TGTGTCAAGTATTTGAGCAGTGGTCCGAGTCTGATCAAGTGGATTTGTAGAACACCTT 360
 121 IleSerGlnMetCysHisTyrglnHisGlyHisLleAsnSerTyLeuLysProMetLeu 140
 361 ATATCCCAAATGTGTCACTACCAGCATGGGCACATCAACTCTTACCTAAACCTATGCTG 420
 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisLleAlaGluAsnIle 160
 421 CAGAGGGATTTCTAACTGCTACCTGACAGCAGCGGGGTCTGGACCATCACTCAGAACATT 480
 161 LeuSerTyLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTy 180
 481 CTGTCACTATGGACGCCAAGTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
 181 ArgValThrSerAspGlyMetLeuTrpLysLeuLlelleGluArgMetValArgThrAsp 200
 541 CGCGTGACGTCACCGCATGCTGTGGAAAAAGCTCATCGAGGAGATGTCAGGACGGAC 600
 201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyVTPGLYGLNTRYLEUPHELYSASNLYS 220
 601 TCTCTGTGGGAGCGCTGGCAGAGCCGAGAGCGGTGGGGACAGTACTTATTCAAAACAAA 660
 221 ProProAspGlyAsnAlaProProAsnSerPheTyArgAlaLeuTyProLysLlelle 240
 661 CTTCTCTGATGAGAAGCTCTCTCCCAACTCTCTTTTATAGAGCGCTTTATCTTAAAAATCATA 720
 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
 721 CAAGACATTGAGACATAGAGTCCAATTTGAGATGTGGCGGACATAGTTTACAGAGAATC 780

261 HisCysArgSerGluThrSerLysGlyValTyrcysLeuGlnTyrcysAspGlnLysile 280
 781 CACTGCCGAGTGAACAAAGTAAAGGGGTCTTACTGTTTACAGTACGACGACCAAGAGATA 840
 281 ValSerGlyLeuArgAspAsnThrIleLysLleTrpAspLysAsnThrLeuGluCysLys 300
 841 GTCAGCGGCTTCGAGAGAACACCATCAAGATCTGGGATAAAGCACACATGGAATGCAAG 900
 301 ArgIleLeuThrGlyHisThrGlySerValLeuLysLeuGlnTyrcysAspGluArgValIle 320
 901 CGGATTTCTACGGGCCACACAGGGCTCCGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
 321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
 961 ATCAGAGGCTCTCTCAGACTCCACCTCAGAGTGTGGATGTAATGTCAGGTGAGATGCTA 1020
 341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
 1021 GACACATTGACTCACCACTGTGAAGCCGTTCTGCACCTGCGCTTCCATTAATGGCATGATG 1080
 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
 1081 GTGACCTGTTCCAAAAGACCGTTCCATCGCTGTGTGGGATATGGCTTCCCAACTGACATC 1140
 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAsp 400
 1141 ACCCTCAGAGGCTGTGTGGGACCCGAGCTGCGGTCAATGTTGTAGACTTTTGATGAC 1200
 401 LysTyrlleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
 1201 NAGTACATGTTTCTGCTCTGGAGATAGAACCATTAAGGTGTGGAAACACAACTACCTGT 1260
 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyLleAlaCysLeuGlnTyrcys 440
 1261 GAACTCGTAAAGACCTTAATTAATGGGACCAAGCGTGGCATCGCTGTTCAGTACAGAC 1320
 441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
 1321 AGCTGTGTGTGAGCGGCTCTCTGACCAACACCATCAGGCTGTGGGACATAGAGTGTGA 1380
 461 AlaCysLeuArgValLeuGluGlyHisGluLeuValArgCysIleArgPheAspAsn 480
 1381 GCATGCTCCGAGTGTGGAGGGCCATGAGGAGTGTGGTACGCTGCATTCGATTTGATAAC 1440
 481 LysArgIleValSerGlyAlaTyrcysGlyLysLleLysValTrpAspLeuValAlaAla 500
 1441 AAAAGATAGTGAAGGAGCGCTATGATGGGAAAAATTAAGTGTGGGATCTTATGGCTGCT 1500
 501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
 1501 TTGGACCGCGTGTCTCCACAGGAGCTCTCTGTCTGGGACACTTGTGGAGCATTTCTGGA 1560
 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
 1561 AGAGTTCCTCGGCTCCAGTTTGTGAATTCAGATTCAGATGTCAGTGTTCATGATGACACA 1620
 541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
 1621 ATTCTCATCTGGAGCTTCTCTGATGATCCAGTGTCTCAGCTGTCTCAGCTGAACCCGCTCCCT 1680
 561 SerArgThrTyThrTyIleSerArg 569
 1681 TCTCGGACATACACCTACATCTCCAGA 1707

RESULT 13

XELSBTRCP 1671 bp mRNA linear VRT 17-SEP-1993
 LOCUS African clawed frog beta-TrCP mRNA, complete cds.
 DEFINITION M98268
 ACCESSION M98268.1 GI:295542
 VERSION beta-transducin repeats.
 KEYWORDS Xenopus laevis (African clawed frog)
 ORGANISM Xenopus laevis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus.
1 (bases 1 to 1671)
Spevak, W., Keiper, B.D., Stratowa, C. and Castanon, M.J.
Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in anaphase are rescued by Xenopus cDNAs encoding N-ras or a protein with beta-transducin repeats
Mol. Cell. Biol. 13 (8), 4953-4966 (1993)
93330289
8393141

Original source text: Xenopus laevis (library: S. cerevisiae expression library of X.laevis oocytes) cDNA to mRNA.

Location/Qualifiers
1..1671
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/cell_type="oocyte"
/tissue_lib="S. cerevisiae expression library of X.laevis oocytes"
1..1671
/gene="beta-TrCp"
48..1604
/gene="beta-TrCp"
/codon_start=1
/protein_id="AA02810.1"
/db_xref="GI:295543"
/translation="MEGFCSLQPTASERDQNRDPKILITKNTLROTQKIANGTSSMTVPQRKLSANVEKEKLVKPFQWSECDQVEVHLISRMCHYQHGHNLYLKPVLRDITFALPARGLDHIAENILSYDLSKLSAEVLCKEYRVYTSIDGMLWKKLIERMVRDLSWRGLAERGRGQYLFKNKPPDGKTPPNSFYRALYPKIIODIETIESNRGGRHSQLRHCRSETSGYVCLQYDDQKIVSGLRDNTIKMDKNTLECKRVLMGHTGSVLCLOYDERVITGSSDSYRVMDVNTGMLNLIHCEAVLHLRFNMGMYTCSKORSIAVMDASATDITLRLVLGHRAAVNVVDQDKYIVSASGRTIKWNTSCFVRILNGHKGILACLOYEDRLVYSGSDNIRLWDLCEGACRLVSGHEELVRCIRPDNKRIVGAYDGKIKWDLVAALDPRAPAGTLCRLTIVHSGRVRLQFDEFQIVSSSHDDTILIWDLNDPGLA"
1671
/gene="beta-TrCp"
/NT 495 a 341 c 424 g 411 t

t Scores:

1. 2.27e-233 Length: 1671
2597.00 Matches: 488
Similarity: 92.87% Conservative: 7
al Similarity: 91.56% Mismatches: 8
tch: 85.60% Indels: 30
5 Gaps: 1

1-168b-2 (1-569) x XELSTRCP (1-1671)

18 SerSerGluArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLys 37
:::|||||
84 GCCTCAGACAGAGACACTGTACAGGATGACCGCTAGAGAGATTATACCGAGAAA 143
38 AsnSerLeuArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrVal 57
:::|||||
144 AACACACTTAGACAG----- 158
58 CysLeuAlaSerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsn 77
159 -----ACAAACTTGCCAAAT 173
78 GlyThrSerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGlu 97
:::|||||
174 GGCACCTCCAGCATGATTGTGCCCAACGACGCAAACTGTGAGCAAAATTACGAGAGAGAA 233
98 LysGluLeuCysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheVal 117
:::|||||
234 AAAGAGCTATGCGTCAAGTATTTTTCAGGACAGTGGTCCGAGTCCGATCAAGTAGAGTTTGT 293

QY 118 GluHisLeuIleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLys 137
Db 294 GAACACCTGATATCTCGAATGTCCACATATCAGCATGGACATATAACACACTTACCTTAAG 353
QY 138 ProMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAla 157
Db 354 CCAATGTTTACAAAGGAGATTTCATACCGCACTPACCAAGTCTCGGGACTCGATCACAATAGCA 413
QY 158 GluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLys 177
Db 414 GAAATATATCTTTCATACCTGGATGCAAGTCAATGTGTCTGCAGAACTGGTATGTAG 473
QY 178 GluTrpTyrArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetVal 197
Db 474 GAGTGTATCGAGTGACCTCAGATGGAATGCTCTGAAAGAACTCATAGAGCGGATGTC 533
QY 198 ArgThrAspSerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTrpLeuPhe 217
Db 534 CGGACAGATTTCTTTTGGAGAGGACTGGCAGAAAGAGGAGTGGGGCAATATTGTGTT 593
QY 218 LysAsnLysProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrPro 237
Db 594 AAAACAAACCTCCAGATGGGAAAGCCACCAAAATTCATCTACAGACCGCTTTACCCA 653
QY 238 LysIleIleGlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeu 257
Db 654 AAAATTATTCAAGACATAGACAAATCGAGTCCAACTGGCGGTGGGAGACACAGCTTA 713
QY 258 GlnArgIleHisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAsp 277
Db 714 CAAAGAAATTCACCTGCGGAGTGAAACAAAGCAAGGGGTATATCTGTCTCAGTACGATGAT 773
QY 278 GlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeu 297
Db 774 CAGAAGATAGTAGTGAAGTGGACTCAGATTAACACCATAGATCTGGGTAAGNATCTTTG 833
QY 298 GluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGlu 317
Db 834 GAGTGCAGCGAGTGTGATGGTCCACTGGGTGAGTCTCTCTGTCGCAATATGATGAG 893
QY 318 ArgValIleIleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGly 337
Db 894 AGAGTAATCATTTACTGGCTCTTCAGACTCTACCGTCCGGGTGTGGACGTGAACACAGGA 953
QY 338 GluMetLeuAsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsn 357
Db 954 GAAATGTTGAACACGCTGATTCACCACTGTGAGGCTGTGCTGCACCTTGAGGTTTAATAT 1013
QY 358 GlyMetMetValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerPro 377
Db 1014 GGCATGATGTCACCTGCTCCAAAGATCGTTCCATTGCAAGTTTGGGACATGGCCTCTGCG 1073
QY 378 ThrAspIleThrLeuArgValLeuValGlyHisArgAlaAlaValAsnValValAsp 397
Db 1074 ACCGATATCACATTACGAGAGATTCTGTGAGCCACCGAGCTGCCGTAAATGTGGTGAC 1133
QY 398 PheAspAspLysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThr 417
Db 1134 TTTGATGACAGATATATAGTCTCATCTGTGTGATCGAACAAATAAGGTTTGGAAACACC 1193
QY 418 SerThrCysGluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGln 437
Db 1194 AGTACATGTGAATTTGTGGGACATTTGAACCGGCCCAAGCGTGGTATTTCATGCTTACAG 1253
QY 438 TyrArgAspArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIle 457
Db 1254 TATCAGATCGGCTTGTGTGTGTGTGTTCTTCTGCAACACATCAGATTGTGGGATATT 1313
QY 458 GluCysGlyAlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArg 477
Db 1314 GAATCGGTGCATGTTTGGGGTTCTGGAGGACATGAAGAACTGGTCCGGTGCATCCGC 1373
QY 478 PheAspAsnLysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeu 497

|||||
1374 TTGTGATACAGAGAAATAGTCAGTGGAGCATATACCGGAAATTAAGTGTGGACCTT 1433
498 ValAlaLeuLeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGlu 517
1434 GTGCTGCTTTGGACCCCGTGGCGCTGCAGGACCTGTGTCTCCGGATCTTGTGGAG 1493

518 HisSerGlyArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHis 537
1494 CATTCGGGCAGAGTTTCCGCGCTGCAGTTTCACGAGTTCACAGATAGTAGAAGCAGCTCTCAC 1553
538 AspAspThrIleLeuIleTyrAspPheLeuAsnAspPro 550
1554 GATGACACTATCCCTCATCTGGATTTCTCAACGCCA 1592

AB033280 2252 bp mRNA linear PRI 14-MAR-2000
ON Homo sapiens BTRCP2 mRNA for F-box and WD-repeats protein
beta-TRCP2 isoform B, complete cds.
AB033280

AB033280.1 GI:7209810
BTRCP2; F-box and WD-repeats protein beta-TRCP2 isoform B.
Homo sapiens (human)

SM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (sites)
Koike,J., Sagara,N., Kirikoshi,H., Takagi,A., Miwa,T., Hirai,M. and
Katoch,M.

Molecular cloning and genomic structure of the betaTRCP2 gene on
chromosome 5q35.1
Biochem. Biophys. Res. Commun. 269 (1), 103-109 (2000)
20160458
D 10694485
2 (bases 1 to 2252)
Katoch,M.

Direct Submission
Submitted (05-OCT-1999) Masaru Katoch, National Cancer Center,
Genetics Division; Tsukiji 5-chome, Chuo-ku, Tokyo 104-0045, Japan
(E-mail:mkatoch@cc.go.jp, Tel:81-3-3542-2511(ex.4402),
Fax:81-3-3541-2685)

Location/Qualifiers
1..2252
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="5"
/map="5q35.1"
/tissue_type="fetal lung"
1..2252
/gene="BTRCP2"
157..1746
/gene="BTRCP2"
/codon_start=1
/product="F-box and WD-repeats protein beta-TRCP2 isoform B"
/protein_id="BA092330.1"
/db_xref="GI:7209811"

/translation="MEPDSVIEDKTIELMNTSVMEQNEDESPPKNTLWQISNGTSVV
IVSRKPSGNYQKEKDIKIFQWSESQVEFVHLIIRMHCHQHGHSYLYKPMW
ORDFTALPEQGLDHAENILSYLDARSLCAELVCKEQRVISEGLMKLTERVWR
TDLWKGLSERGMDVLFKRPDTGPPNSFYSLPKLIQDIETISNWRGGRHLQ
RIQRENSKGYCLQVDDEKIIISGLRNSIKIWDKISLECLVLTGHTGSVLQCLQD
ERVIVTSSDSTVRVMDNTGEVNTLIHNEAVLHLEFNSGLMTVCSDRSIAVMDM
ASADITLRLVYVGHRAVNVVDKDIYVSAGDRTIKWSTSTCFEFTLNGHRKK
IACQYRDLRVSGSSONTIRLWIEGACILRVLEGEELVRICRPNKRVISGVDG
KIKYWDLQALDPRAPASTLCLRLVHESGRVFLQDFEQIISSSDDHLLIWDPLN
VPPSAQNETRFSRTYTVISR"

NT 598 a 501 c 558 g 595 t

it Scores:

Pred. No.:	9,69e-215	Length:	2252
Score:	2400.00	Matches:	451
Percent Similarity:	86.82%	Conservative:	43
Best Local Similarity:	79.26%	Mismatches:	35
Query Match:	79.10%	Indels:	40
DB:	9	Gaps:	4
US-09-601-168B-2 (1-569) x AB033280 (1-2252)			
QY 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20			
Db 157 ATGGAGCCC---GACTCGTGATTGGAGCAGACCACCTCAGCTCATGAACACTTCAGTT 213			
QY 21 ArgGluAspCysAsnAsnGlyGluProProArgHisIleIleProGluLysAsnSerLeu 40			
Db 214 ATGGAAGATCAAAATGAAGATGAGTCCCAAG-----AAAAATACTCTT 258			
QY 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60			
Db 259 TGGCAG----- 264			
QY 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80			
Db 265 -----ATAAGTAATGGAACATCA 282			
QY 81 SerMetIleValProGlyGlnArgLysLeuSerLysLeuLysGluLysGluLeu 100			
Db 283 TCTGTGATCGTCTCCAGAAAGAGGCCATCAGAAAGAACTATCAAAAAGAAAAGACTTG 342			
QY 101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120			
Db 343 TGTATTAAATATTTTGACCAAGTGCTGAATCAGATCAAGTGGAAATTTGTGAACATCTT 402			
QY 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140			
Db 403 ATTTCCAGAAATGTGTCTATTCAGCATGGACATATTAACTCTTACCTCAAGCCCATGTTG 462			
QY 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160			
Db 463 CAGCGGGACTTTTATACCGCTTTACCGAGCAAGGCTTAGATCATAGCAGAAAACATT 522			
QY 161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 180			
Db 523 CTTTGTACTGGATGCCAGGTCCTCTGTGCGACGAGAGCTGGTATGTAAAGAAATGGCAG 582			
QY 181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200			
Db 583 CGAGTGATCTCAGAAAGGAATGCTTTGGAAGAAGCTGATTGAACGAATGGTACGCACATGAT 642			
QY 201 SerLeuTyrArgGlyLeuAlaGluArgGlyTyrGlyGlnTyrLeuPheLysAsnLys 220			
Db 643 CCCCTATGGAAGAGACTTTCAGAAAGAAGAGGGTGGGATCAGTACCTGTTTAAAAACAGA 702			
QY 221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240			
Db 703 CCCACAGATGGC-----CCTCCAAATTCATTTTATAGTCAATATATATCCCAAGATTATC 756			
QY 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260			
Db 757 CAGGATATAGACACTATAGAATCTTAACCTGGCGGTGGACGACACAACTTGCAGAGGATT 816			
QY 261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280			
Db 817 CAGTCCCGCTCTGAAAATAGTAAGTGTCTACTCTTTTACAGTACGATGATGAATAAATT 876			
QY 281 ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300			
Db 877 ATCAGTGGCTACGAGATAATTCATTATTAAGATAGGATAAACCAGCGCTGGAATGTTG 936			
QY 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320			
Db 937 AAAGTGTTAAACAGGACACACAGGCTCTGTCTCTCTGTCAGTATGATGAGCGTGTCTATT 996			

321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
:::|||||
997 GTAACTGGCTCTTACAGATTCTACCGGTGAGATGGTGGGATGGAACACGGGTGAAGTCTT 1056
341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
1057 AACACATTGATCCACCACCAATGAGGCTGTATTGGACATTACGCTTCAGCATGGACTGATG 1116
361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
1117 GTGACCTGTTCAGAGACCGCTCCATTCCTGTGTGGGACATGGCTTCTGGCAGCGACATC 1176
381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
1177 ACTTTACGCGGTCTCTGTGTGGCCACGCGGCTGCCGTCAATGATGATAGACTTTGACGAC 1236
401 LysTyrlleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
1237 AAGTACATCGTGTCTGCCCTCTGGTGACAGGACCATCAAGTCTGGAGCAGCAGCACCTGT 1296
421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyArgAsp 440
1297 GAATTTGTTCGTACTCTCAATGGGCACCAAGCGGGCATTCGCTGTCTCCAGTACAGGGAT 1356
441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
1357 CGCTGTGTTGTAGTGGATCATCAGATATACCATAGCTCTGGGATATTGAATGTGGT 1416
461 AlaCysLeuArgValLeuGluGlyHisGluLeuValArgCysIleArgPheAspAsn 480
1417 GCCTGTTTAAAGAGTCTTAGAGGACATCAAGAAATTTGGTCCGATGCATCCGGTTTGATAAC 1476
481 LysArgIleValSerGlyValaTrpAspGlyLysIleLysValTrpAspLeuValAla 500
1477 AAGAGATTGTCACTGGGCGCTTATGATGGGAAATTAAGTTTGGACTTCCAGACTGCT 1536
501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
1537 CTGACCTCTGAGCGCCAGCAGCAGCATTTGTGTTTGGCAGCATTTGGTGGACATCTCTGA 1596
521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerSerHisAspThr 540
1597 CGTGTGTTTCGGCTCCAGTTTCAATGAGTTTCAATGAGTTTCAATGAGTTTCAATGAGTT 1656
541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
1657 ATTTTGAATGGGATTTCTTAATGTGCTCCAGTCCGACATGAGACCCGTTCTCCC 1716
561 SerArgThrTyThrTyrlleSerArg 569
1717 TCCAGAACATACATCTACATCTCTAGA 1743

5

BC026213 4395 bp mRNA linear PRI 08-APR-2002
ON Homo sapiens, F-box and WD-40 domain protein 1B, clone MGC:21122
IMAGE:4419029, mRNA, complete cds.
N BC026213
MGC:21122
MGC:21122
Homo sapiens (human)
SM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
E 1 (bases 1 to 4395)
S Strausberg, R.
Direct Submission
L Submitted (02-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk

Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Prepared by: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 27 Row: i Column: 22.

FEATURES

source

1..4395
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocusID:23291"
/db_xref="taxon:9606"
/clone="MGC:21122 IMAGE:4419029"
/issue_type="Lymph, lymphoma"
/clone_lib="NIH_MGC_85"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
372..1961
/codon_start=1
/product="F-box and WD-40 domain protein 1B"
/protein_id="AAH26213.1"
/db_xref="GI:20070728"
/translation="MEPDSVIEDTIELMNTSVMEDQNEDESPKKNTLWQISNTSSV
IVSRKREGNYQKEKDLCKYFQWSESDQVEFVHLISRMCHYQHGHSYKPLM
ORDFTALPEOGLDHAENILSYLDASLCAELVCKEQRVISEGMLKKLIERMVR
TDPLWKGLSERGMDOVLFRNPDTGPNSEFVSLYPKIIOIETIESNWRGRRNLQ
RIQSESNKGVYCLQYDDEKILISGLRDSIKINDKTSLECLVLTGHTGSLVCLQYD
ERVITGSSDSTVRVMDVNTGEVLTLLHNEAVLHURFVSNGLMVTCSKRSIAWDM
ASATDITLRRVLVGHRAAVNYDFDDKTVISAGDRTIKVNSTCTCFVRLNGLKRG
IACQLDRLLVVGSSDNTIRLWIECCAGLVEGLHEELVRCIRFONKRTIVSAYDG
KIKWDLQALDPRAPASTCLRLTVHSGRVFLQPDFQIISSHDDTLILWDFLN
VPPSAQNETRSPSTTYISR"

BASE COUNT 1131 a 928 c 1039 g 1297 t

ORIGIN

Alignment Scores:
Pred. No.: 2,42e-214 Length: 4395
Score: 2400.00 Matches: 451
Percent Similarity: 86.82% Conservative: 43
Best Local Similarity: 79.26% Mismatches: 35
Query Match: 79.10% Indels: 40
DB: 9 Gaps: 4

US-09-601-168B-2 (1-569) x BC026213 (1-4395)

QY 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
Db 372 ATGGAGCCC---GATTCGGTGAATGAGGACAGACCATCGAGCTCATGAACATTCAGTT 428
QY 21 ArgGluAspCysAsnAsnGlyGluProArgLysIleIleProGluLysAsnSerLeu 40
Db 429 ATGGAAGATCAAAATGAAGATGATGCCCAAG-----AAAAATACTCTT 473
QY 41 ArgGlnThrTyArgSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
Db 474 TGSCAG-----
QY 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
Db 480 -----ATAAGTAATGAACATCA 497
QY 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrgLysGluLysGluLeu 100
Db 498 TCTGTGATCGTCTCCAGAAAGAGGCCATCAGAGGAACCTATCAAAAGAAAAAGACTTG 557


```
101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
    ::::::::::::::::::::
558 TGTATTAAATATTTGACCAGTGGTCTGAATCAGATCAAGTGGAAATTTGTGGAACATCTT 617
    ::::::::::::::::::::
121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
    ::::::::::::::::::::
618 ATTTCAGCAATGTGCTATTATCAGATGAGACATATTAACTCTTACCTGGAAGCCCATGTTG 677
    ::::::::::::::::::::
141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
    ::::::::::::::::::::
678 CAGCGGGACTTTATTACCGCTTTTACAGAGCAAGGCTTAGATCAGATGAGGAAACAATT 737
    ::::::::::::::::::::
161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 180
    ::::::::::::::::::::
738 CTTTCGTACTCGATGCCAGGCTCTCTGTGTGCAGCAGAGCTGGTATGTAAAGAAATGGCAG 797
    ::::::::::::::::::::
181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
    ::::::::::::::::::::
798 CGAGTGATCTCAGAGGAAATGCTTTGGAAAGAAAGCTGATTAACGAATGGTACGCACCTGAT 857
    ::::::::::::::::::::
201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTyrGlyGlnTyrLeuPheLysAsnLys 220
    ::::::::::::::::::::
858 CCCATATCGAAAGGACTTTCAGAAAGAGAGAGGCTGGGATCAGTACCTCTTTAAAAACAGA 917
    ::::::::::::::::::::
221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
    ::::::::::::::::::::
918 CCCACAGATGGC-----CCTCCAAATTCATTTATAGGTCAATTATACCCAAAGATTATC 971
    ::::::::::::::::::::
241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
    ::::::::::::::::::::
972 CAGGATATAGAGACTATAGAATCTAACTGGCGGTGTGCAGCAGACACAACCTTCAGAGGATT 1031
    ::::::::::::::::::::
261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
    ::::::::::::::::::::
1032 CAGTGGCGCTCTGAAATATGTAAGAGTGCTACTGTTTACAGTACGATGATGATGATGATG 1091
    ::::::::::::::::::::
281 ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300
    ::::::::::::::::::::
1092 ATCAGTGGCTACGAGATTAATTCATTATAGATATGGGATAAACACAGCTGGAAATGTTG 1151
    ::::::::::::::::::::
301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
    ::::::::::::::::::::
1152 AAAGTGTTAAACAGGACACACAGGCTCTGTCTCTGTCTGAGTATGATGAGCGTGTCAAT 1211
    ::::::::::::::::::::
321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
    ::::::::::::::::::::
1212 GTAACGTGGCTCTTCAGATCTACGGTGAGAGTGGGATGTGAACCGGGTGAAGTCTTT 1271
    ::::::::::::::::::::
341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
    ::::::::::::::::::::
1272 AACACATTGATCCACCACAATGAGGCTGTATTGCATTTACGCTTCAGCAATGGACTGATG 1331
    ::::::::::::::::::::
361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
    ::::::::::::::::::::
1332 GTGACCTGTTCCAAAGGACCGCTCCATGTGTGTGGGACATGGCTTCTGCGACCGACATC 1391
    ::::::::::::::::::::
381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
    ::::::::::::::::::::
1392 ACTTTACGCGGTGCTCTGTGTGGCCACCGGGGTGGCTCAATGTAGTACTTTGACGAC 1451
    ::::::::::::::::::::
401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
    ::::::::::::::::::::
1452 AAGTATCATGCTGTCTGCTCTGGTACAGGACCATCAAAAGTCTGGAGCAGCAGCACCTGT 1511
    ::::::::::::::::::::
421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
    ::::::::::::::::::::
1512 GAATTTGTTCGTACTCTCAATATGGGCACAGCGGGGCAATTTGCCCTGTCTCAGTACAGGAT 1571
    ::::::::::::::::::::
441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
    ::::::::::::::::::::
1572 CGCTGGTGTATTAGTGAATCATCAGATATATACCATTAGGCTCTGGGATATTGAATGTGCT 1631
    ::::::::::::::::::::
461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
```

```
Db 1632 GCCTGTTTAAAGAGTCCTAGAGGGACATGAAGAAATTGGTCCGATGCATCCGGTTTGATAAC 1691
    ::::::::::::::::::::
QY 481 LysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTyrAspLeuValAlaAla 500
    ::::::::::::::::::::
Db 1692 AAGAGGATTTGTCAGTGGGGCCCTATATGGGAAAATTAAGATTTGGGACTTGCAAGCTGCT 1751
    ::::::::::::::::::::
QY 501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
    ::::::::::::::::::::
Db 1752 CTTGACCTCGAGCCCCAGCAAGCACATTGTGTTGCGCACATTGGTGGAAACATTCTGGA 1811
    ::::::::::::::::::::
QY 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerSerHisAspAspThr 540
    ::::::::::::::::::::
Db 1812 CGTGTGTTTCGGCTCCAGTTTTGATGAGTTTCAGATCATCAGCAGCTCCCATGATGACACT 1871
    ::::::::::::::::::::
QY 541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
    ::::::::::::::::::::
Db 1872 ATTTTGATTTGGGATTTCTTAAATGTGCTCCAGTGCCGAGATGAGACCCGTTCTCCC 1931
    ::::::::::::::::::::
QY 561 SerArgThrTyrThrTyrIleSerArg 569
    ::::::::::::::::::::
Db 1932 TCCAGAACATACACTTACATCTCTAGA 1958
    ::::::::::::::::::::
```

Search completed: October 22, 2003, 16:26:46
Job time : 5030 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

ein - nucleic search, using frame_plus_p2n model

October 22, 2003, 14:49:27 ; Search time 2665 Seconds
(without alignments)
5189.211 Million cell updates/sec

US-09-601-168B-2
score: 3034
e: 1 MPDAEVLQEKALFPMNSSE.....PAAQAEPPSPSRYYTIVISR 569

table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

d: 22781392 seqs, 12152238056 residues

umber of hits satisfying chosen parameters: 45562784

DB seq length: 0

DB seq length: 2000000000

rocessing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

line parameters:

framed_p2n.model -DRV=xlp
21/USPTO.spool_p/US09601168/runat_22102003_100546_17791/app_query.fasta_1.711
-QMT=fastap -SUFFIX=1 -MATRIX=blosum62 -TRANSHUMAN40.cdi -LIST=45
bits -START=1 -END=1 -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
EN=200 -THR SCORE=pct -THR MAX=100 -MINLEN=0 -MAXLEN=200000000
-pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
303601168 @CGN_1.1 2874 @runat_22102003_100546_17791 -NCPU=6 -ICPU=3
>LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
4EOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
r=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

3 : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2951	97.3	2897	11	AK032221	AK032221 Mus muscu
2	2940	96.9	2970	11	AK083257	AK083257 Mus muscu
3	2686	88.5	4502	11	AK052317	AK052317 Mus muscu
4	2395	78.9	3824	11	AK041532	AK041532 Mus muscu
5	1561	51.5	1136	14	CD500760	CD500760 CDA46-H04
6	1407.5	46.4	924	14	CA985560	CA985560 AGENCOURT
7	1390	45.8	895	9	AL898993	AL898993 AL898993
8	1353	44.6	966	14	CA980902	CA980902 AGENCOURT
9	1350	44.5	776	14	CB244526	CB244526 UI-M-FY0-
10	1348	44.4	810	12	BM944304	BM944304 UI-M-EH0p
11	1332	43.9	778	13	BQ444188	BQ444188 UI-M-EX0-
12	1323.5	43.6	1002	14	CA971789	CA971789 AGENCOURT
13	1320.5	43.5	935	14	CA975032	CA975032 AGENCOURT
14	1310	43.2	850	14	CA327271	CA327271 UI-M-FY0-
15	1306	43.0	744	14	CB520936	CB520936 UI-M-GH0-
16	1298.5	42.8	850	13	BQ771102	BQ771102 UI-M-FI0-
17	1298.5	42.8	1201	9	AL556068	AL556068 AL556068
18	1290	42.5	749	14	CB246650	CB246650 UI-M-FI0-
19	1270	41.9	806	13	BU750387	BU750387 CH3#032 H
20	1241	40.9	796	14	CA348002	CA348002 679219 NC
21	1238	40.8	791	14	CA320826	CA320826 UI-M-FW0-
22	1238	40.8	827	14	CB524867	CB524867 UI-M-FY0-
23	1227	40.4	804	14	CB790902	CB790902 AGENCOURT
24	1226.5	40.4	972	13	BQ940482	BQ940482 AGENCOURT
25	1225	40.4	699	14	CB246901	CB246901 UI-M-FI0-
26	1222.5	40.3	876	13	BU170327	BU170327 AGENCOURT
27	1222	40.3	732	12	B1088274	B1088274 602852735
28	1218	40.1	888	12	B1455545	B1455545 603173974
29	1214	40.0	846	9	AL879775	AL879775 AL879775
30	1184	39.0	1081	12	BM542586	BM542586 AGENCOURT
31	1184	39.0	1201	9	AL531016	AL531016 AL531016
32	1181	38.9	870	13	BQ463637	BQ463637 603772881
33	1177	38.8	851	13	BQ425358	BQ425358 AGENCOURT
34	1174	38.7	963	13	BX349157	BX349157 BX349157
35	1169	38.5	565	9	AL642623	AL642623 AL642623
36	1169	38.5	758	13	BUG11165	BUG11165 UI-M-FI0-
37	1168	38.5	523	13	BQ913488	BQ913488 AGENCOURT
38	1166	38.4	973	12	BM465280	BM465280 AGENCOURT
39	1157	38.1	667	9	AL882125	AL882125 AL882125
40	1154	38.0	712	14	CA513313	CA513313 UI-R-FJ0-
41	1149	37.9	985	13	BX419479	BX419479 BX419479
42	1146	37.8	752	9	AU002962	AU002962 AU002962
43	1143	37.7	779	14	CD350839	CD350839 UI-M-GI0-
44	1143	37.7	699	13	BUG13203	BUG13203 UI-M-EW0-
45	1141	37.6	761	14	CA345696	CA345696 676407 NC

ALIGNMENTS

RESULT 1
AK032221
LOCUS
DEFINITION
AK032221 2897 bp mRNA linear HTC 05-DEC-2002
Mus musculus adult male olfactory brain cDNA, RIKEN full-length
enriched library, clone:6430504E23 product:beta-transducin repeat
containing protein, full insert sequence.
ACCESSION
VERSION AK032221.1 GI:26082733
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

CE 1 Carninci, P. and Hayashizaki, Y.
 RS High-efficiency full-length cDNA cloning
 AL Meth. Enzymol. 303, 19-44 (1999)
 NE 99279253
 ED 10349636
 CE 2
 RS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 AL Genome Res. 10 (10), 1617-1630 (2000)
 NE 20493374
 ED 11042159
 CE 3
 RS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Mishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (KISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 AL Genome Res. 10 (11), 1757-1771 (2000)
 NE 20530913
 ED 11076861
 CE 4
 RS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batalow, S., Casavant, T.,
 Fleischmann, W., Gaasterland, T., Gassi, C., King, B., Kochiwa, H.,
 Kuehl, P., Lewis, S., Matsumoto, Y., Nikaide, I., Pesole, G.,
 Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
 Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
 Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
 Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L.,
 Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
 and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 AL Nature 409 (6821), 685-690 (2001)
 NE 21085660
 ED 11217851
 CE 5
 RS The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 AL Nature 420, 563-573 (2002)
 CE 6 (bases 1 to 2897)
 RS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/
 URL: http://Location/Qualifiers

source
 1. .2897
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM_DB:6430504E23"
 /db_xref="taxon:10090"
 /clone="6430504E23"
 /sex="male"
 /tissue_type="olfactory brain"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 misc_feature
 1. .2897
 /note="beta-transducin repeat containing protein
 (MGD|MG1:1338871, GB|NM_009771, evidence: BLASTN, 100%,
 matches=1502)"

BASE COUNT 720 a 736 c 771 g 670 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.37e-314 Length: 2897
 Score: 2951.00 Matches: 561
 Percent Similarity: 93.07% Conservative: 3
 Best Local Similarity: 92.57% Mismatches: 5
 Query Match: 97.26% Indels: 37
 DB: 11 Gaps: 1

US-09-601-168B-2 (1-569) x AK032221 (1-2897)

Qy 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMet----- 16
 Db 68 ATGGACCCGGCAGAGCGGTGCTGCAGAGAAAGCGCTTAAGTTATGTGCTCTATGCC 127
 Qy 16 ----- 16
 Db 128 AGGTCTCTGTGGCTGGGCTGCTCCAGCCTGGCGGACAGCATGCCCTTCGCTGCATGCC 187
 Qy 17 -----AsnSerSerGluArgGluAspCy 24
 Db 188 TATAACCCAGGCACCTGGCGCACTCAAGCTTTCCAGAAATTCCTCAGAGAGAGAACTG 247
 Qy 24 sAsnAsnGluGluProProArgLysIleIleProGluLysAsnSerLeuArgGlnThrTy 44
 Db 248 TATAATGGCGAACCCCTTAGAAGTAATAACAGAGAAAGATTCACCTAGACAGACTTA 307
 Qy 44 rAsnSerCysAlaAArgLeuCysLeuAsnGlnGluThrValCysLeuAlaSerThrAlaMe 64
 Db 308 CAACAGCTGTGGCAGGCTTTGCATAAACCAGAGACAGTATGCTTAACAGCAGCTGCTAT 367
 Qy 64 tlyThrGluAenCysValAlaLysThrLysLeuAlaAenGlyThrSerSerMetIleVa 84
 Db 368 GAAGACTGAAATTTGTGTGGCCAAAGCCAAATTCGCAATGGCACTTCCAGCATGATTGT 427
 Qy 84 lProLysGlnAArgLysLeuSerAlaSerTyrcGluLysGluLysGluLeuCysVallysty 104
 Db 428 GCCCAGCAGCGGAAACTCTCAGCAAGCTATGAGAGAGAAAGGAGAGCTGTGTCAAGTA 487
 Qy 104 rPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeuIleSerGlnMe 124
 Db 488 TTTTGGCAGTGGTCAGACTCTGATCACTGGAATTTGTAGAACACCTTATATCCAAAT 547

```

124 tCysHisTyrGlnHisGlyHisIleAsnSerTyrIleuLysProMetIleuGlnArgAspPh 144
548 GTGTCACTACCAAGCATGGGCACATCAACTCTCTACCTAAACCTATGCTGAGAGGGATT 607
144 eileThrAlaLeuProAlaAArgGlyLeuAspHisIleAlaGluAsnIleLeuSerTyrLe 164
608 CATAACTGCATCGCAGACAGGGGCTCTGGACCATCGCTGAGAACATCTGTCTATACTT 667
164 uAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyrArgValThrSe 184
668 GGACGCCAAGTCACTGTGTGCTGTGCTGAGCTGCTGCAAGGAATGTTACCGCGTGAAGTC 727
184 rAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAspSerLeuTrpAr 204
728 GGACGCATGCTGTGTAAGAAACCTCATCGAGGAGTGTGAGGCGGACTCTGTGTGGCG 787
204 gGlyLeuAlaGluAArgGlyTrpGlyGlnTyrLeuPheLysAsnLysProProAspGl 224
788 AGGCTTGGCAGAGCGCAGAGGCTGGGACAGTACTTATCAAAACAAACCTCTCTGATGA 847
224 yAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleGl 244
848 GAACGCTCTCTCCCACTCTCTTTATAGAGCGCTTTATCTCTAAATCATACAAGACATTGA 907
244 uThrIleGluSerAsnTrpAtGysGlyAArgHisSerLeuGlnArgIleHisCysArgSe 264
908 GACATAGAGTCAATTGGAGATGTGGGCGACATAGTTTACAGAGATCCACTCCCGGAG 967
264 rGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLe 284
968 TGAACAACAGTAAAGGGGTTTACTGTGTTACAGTACGACGACAGACAGATAGTCAGGGGCT 1027
284 uArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLysArgIleLeuTh 304
1028 TGAGACAAACCATCAAGATCTGGATTAAGACACATCTGGATGCAAGCGGATCTTCAC 1087
304 rGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIleIleThrGlySe 324
1088 GGGCCACACGGGCTCGTCTGCTGCTGCTGAGTACGATGAGAGGAGTATCATCAGAGGCTC 1147
324 rSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIl 344
1148 CTCAGACTCCACCGTCAGAGTCGGGATGTAATTCAGCTGAGATGCTTAACACATATGAT 1207
344 eHisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSe 364
1208 TCACCACTGTGAAGCGTTCTGCACCTGGCTTCAATAATGCGATGATGGTACCTGTTTC 1267
364 rLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgAr 384
1268 CAAGACCGTTCATCGCTGTGTGGGATATGGCTTCCCACTGACATCACCCCTCAGGAG 1327
384 qValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspLysTyrIleVa 404
1328 GTGTGCTGTGGGACACCGAGCTGGGCTCAATGTTGATGCTTGTAGCTTGTGATGATCATCGT 1387
404 lSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCysGluPheValAr 424
1388 TTCTGCTCTGGAGATAGAACCAATAAAGGTGTGGAACACACAGTACCTGTGAATTCGTAA 1447
424 gThrLeuAsnGlyHisArgGlyIleAlaCysLeuGlnTyrArgAspArgLeuValVa 444
1448 GACCCCTAAATGGGCACCAAGCGTGGCATCGCTGTTGCGATGACAGAGACGCTGGTGGT 1507
444 lSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGlyAlaCysLeuAr 464
1508 GAGCGGCTCTCTGCACAAACCACTACGCTGTGGGCATAGAGTGTGGAGCATGCTCGC 1567
464 gValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsnLysArgIleVa 484
1568 AGTGTGGAGGGCCATGAGGAGTTGCTACGCTGCATTCGATTTGATTAACAAAGATAGT 1627

```

484 lSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAlaLeuAspProAr 504
1628 GAGCGGACCCCTATGATGGGAAATTAAGTGTGGATCTTATGCTGCTTTGGACCGCG 1687
504 gAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGlyArgValPheAr 524
1688 TGCTCCACAGGGACTCTCTGCTCGGACACTTGTGGAGCATCTCGAAGAGATTTCCG 1747
524 gLeuGlnPheAspGluPheGlnIleValSerSerSerHisAspAspThrIleLeuIleTr 544
1748 CTCCTCAGTTTATGAATTCAGATTTGTCAGTAGTTCATGATGACACAAATTCATCTG 1807
544 pAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerProSerArgThrTy 564
1808 GGACTTCTGATGATCCAGTGTCTACGCTGACCGCGCTCCCTTCTCGGACATA 1867
564 rThrTyrIleSerArg 569
1868 CACCTACATCTCCAGA 1883

RESULT 2
AK083257 2970 bp mRNA linear HTC 05-DEC-2002
LOCUS Mus musculus adult male hippocampus cDNA, RIKEN full-length
DEFINITION enriched library, clone:C630031K01 product:beta-transducin repeat
containing protein, full insert sequence.
ACCESSION AK083257
VERSION AK083257.1 GI:26101148
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Rashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujisawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multiplexed sequence
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamakawa, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Staab, P., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,

Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzei, J., Mombert, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection

Nature 409 (6821), 685-690 (2001)

21085660

11217851

5

CE

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

RS

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 2970)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saichou, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-APR-2002) Yoshihide Hayashizaki. The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,

Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

Location/Qualifiers

1..2970

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="PANTOM:DB:C630031K01"

/db_xref="taxon:10090"

/clone="C630031K01"

/sex="male"

/tissue_type="hippocampus"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

81..1900

/note="beta-transducin repeat containing protein

(NGF|MG1:1338871, GB|NM_009771, evidence: BLASTN, 100%,

match=1502)

putative"

sc_feature

UNT 739 a 758 c 788 g 685 t

nt Scores:

O.: 2.35e-313 Length: 2970

2940.00 Matches: 560

92.92% Conservative: 4

92.26% Mismatches: 5

96.90% Indels: 38

11 Gaps: 1

US-09-601-168B-2 (1-569) x AK083257 (1-2970)

Qy 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMet----- 16

Db 81 ATGACCCCGCAGCGGGTGTGTCAGAGAGAAAGCGCTTAAGTTTATGTCTTATGCC 140

Qy 16 ----- 16

Db 141 AGGTCTCTGTGGCTGGGTCTCTCAGCCTGGCGGACAGCATGCTTCTCGCTGGCATGCCTG 200

Qy 17 -----AsnSerSerGluArgGluAspCys 24

Db 201 TATAACCCAGGAGACTGGGCACCTCAAGCTTCCAGAAATTCCTCAGAGAGAGAACTGT 260

Qy 25 AsnAsnGlyGluProArgLysIlelleProGluLysAsnSerLeuArgGlnThrTyr 44

Db 261 AATAATGGCGAACCCTTAGGAAGATAAATACCAGAGAAATTCCTTAGACAGACTTAC 320

Qy 45 AsnSerCysAla-ArgLeuCysLeuAsnGlnGluThrValCysLeuAlaSerThrAlaMe 64

Db 321 AACAGCTGTGCCAGGCTTTGCATAAACCAAGAGACAGATATGTCTAACAGCACTGCTAT 380

Qy 64 tLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSerSerMetIleVa 84

Db 381 GAAGACTGAAATTTGTGGCCAAAGCCAAATTCCTTCCAGCAATGATTGT 440

Qy 84 lProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLysValLysTy 104

Db 441 GCCCAAGCAGCGGAAACTCTCAGCAAGCTATGAGAGAGAAAGAGCTGTGTGTCAAGTA 500

Qy 104 rPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisIleSerGlnMe 124

Db 501 TTTTGAGCAGTGTGTGAGAGTCTGATCAAGTGGAAATTTAGAACACCTTATATCCCAAT 560

Qy 124 tCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeuGlnArgAspPh 144

Db 561 GTGTCACTACCAGCATGGGCACATCACTCTACCTAAACCTATGCTGCAGAGGGATT 620

Qy 144 elleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIleLeuSerTyrIe 164

Db 621 CATAACTGCACCTCCAGCAGCGGGTGTGGACACATCGCTGAGAACATTCGTCTCATAC 680

Qy 164 uAspAlaLysSerLeuCysAlaLalaGluLeuValCysLysGluTrpTyrArgValThrSe 194

Db 681 GGACGCCAAGTCACTGTGTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 740

Qy 184 rAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAspSerLeuTrpAr 204

Db 741 GGACGCCATGCTGTGGAAAAAGCTCATCGAGAGAGTGGTTCAGGACGGACTCTCTGTGGCG 800

Qy 204 gGlyLeuAlaGluArgGlyTrpGlyGlnTyrLeuPheLysAsnLysProAspGly 224

Db 801 AGGCTTGGCAGCGCAGAGGCTGGGACAGTACTTATTCAAAAACAAACCTCTCTGATGA 860

Qy 224 YAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleG 244

Db 861 GAACGCTCTCTCCCACTCTCTTTTATAGAGCGCTTTATCTTAAATCATACAGACATTGA 920

Qy 244 uThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSe 264

Db 921 GACAAATAGATCCCAATTTGGAGATGTGGCGACATAGTTTACAGAGAATCCACTGCCGGAG 980

Qy 264 rGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLe 284

Db 981 TGAACAAGTAAAGGGGTTTACTGTCTTACAGTACGACGACACAGAGATAGTCAGCGGCT 1040

Qy 284 uArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLysArgIleLeuTh 304

Db 1041 TCGAGACAAACCATCAAGATCTGGGATAAAGACACCTGGAATGCAAGCGGATTCTCAC 1100

Qy 304 rGlyHisThrGlySerVal-LeuCysLeuGlnTyrAspGluArgValIleIleThrGlyS 324

1101	GGGCCACACGGGCTCCGTCCTCTGTCTGTCAGTACAGTACGAGGGGTGATCATCAGGCT	116
324	erSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuI	344
1161	CCTCAGACTCCACCGCTCAGAGTGTGGGATGTAATGACAGTGAGATGCTAAACACATTGA	1220
344	leHisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetValThrCysS	364
1221	TTTCAACACTGTGAAGCGGTTCTGCACCTGGCGTTCAATTAATGGCATGATGGTGACCTGTT	1280
364	erLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgA	384
1281	CCAAAGACCGTTCATCGCTGTGTGGGATTGGCTTCCCACTGCATCATCCCTCAGGA	1340
384	rgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspLysTyrIleV	404
1341	GGGTGCTGTGGGACACCGAGCTCGGTCAATGTTGTAGACTTTGATGACAAGTACATCG	1400
404	alSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCysGluPheValA	424
1401	TTTCTGCTCTGGAGATAGAACCAATAAGGTGTGGAAACAAGTACTCTGTGAAATTCGTAA	1460
424	rgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAspArgLeuValV	444
1461	GGACCTTAATGGGCACAAAGCTGGCATCGCCTGTTTGAGTACAGACAGGCTGGTGG	1520
444	alSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGlyAlaCysLeuA	464
1521	TGAGGGCTCCTCTGCACAAACCATCAGGCTGTGGACATAGAGTGTGGAGCATGCCTGC	1580
464	rgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspLysArgIleV	484
1581	GAGTGTGTGGAGGCCCATGAGGAGTGTGTAGCTGCATTCGATTGATTAACAAAAGGATAG	1640
484	alSerGlyAlaTrpAspGlyLysIleLysValTrpAspLeuValAlaLeuAspProA	504
1641	TGAGGGGAGCCTATGATGGGAAATTAAGGTGTGGATCTTATGGCTTGTGGACCCGC	1700
504	rgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGlyArgValPheA	524
1701	GTGCTCCAGCAGGACTCTCTGTCTGGGACACTTGTGGAGCATCTCGGAGAGTTTTC	1760
524	rgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThrIleLeuIleT	544
1761	GGCTCCAGTTTGATGAATTCAGATTGTCAGTAGTTTCATGATGACACAAATTCATCT	1820
544	rpAspPheLeuAsnAspProAlaAlaGlnAlaGluProArgSerProSerArgThrT	564
1821	GGGACTTCTGATGATCCAGCTGCTCACGCTCAACCGCCCGCTCCCTTCGGACAT	1880
564	yrThrTyxIleSerArg	569
1881	ACACCTACATCTCCAGA	1897
AK052317	4502 bp	mRNA
ON	Mus musculus 13 days embryo heart cDNA, RIKEN full-length enriched library, clone:D330027D11 product:beta-transducin repeat containing protein, full insert sequence.	
N	AK052317	
	AK052317.1	GI:26095121
	HTC; CAP trapper.	
	Mus musculus (house mouse)	
SM	Mus musculus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
E	1	
S	Carninci, P. and Hayashizaki, Y.	
	High-efficiency full-length cDNA cloning	
L	Meth. Enzymol. 303, 19-44 (1999)	
E	99279253	
	10349636	

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes *Genome Res.* 10 (10), 1617-1630 (2000)
20499374
11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer *Genome Res.* 10 (11), 1757-1771 (2000)
20530913
11076861

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yananaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischnmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Kuehnbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y. Functional annotation of a full-length mouse cDNA collection *Nature* 409 (6821), 685-690 (2001)
20185660
11217851

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs *Nature* 420, 563-573 (2002)
12466851

6 (bases 1 to 4502)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katch, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurhara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Soqabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-resesg.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>

URL: <http://fantom.gsc.riken.go.jp/>

Location/Qualifiers

1. 4502

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="PANTOM_DB:D330027D11"

/db_xref="taxon:10090"

/clone="D330027D11"

/tissue_type="heart"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stages="13 days embryo"

1. 4502

/note="beta-transducin repeat containing protein (MGD|GI:1338871, GB|NM_009771, evidence: BLASTN, 100%, match=1502)"

JNT 1284 a 998 c 1126 g 1094 t

nt Scores:

5.72e-285 Length: 4502

2686.00 Matches: 509

90.46% Conservative: 3

89.93% Mismatches: 4

88.53% Indels: 50

11 Gaps: 1

01-168B-2 (1-569) x AK052317 (1-4502)

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMet----- 16

88 ATGGACCCGGCAGAGCGGTGCTGAGGAGAAAGCGCTTAAGTTATGGTGTCAAGTCCG 147

16 ----- 16

148 GCTGGTGATATTGAGGAATAATAAAGTCTCTATGCCAGGTCTCTGGCTGGC 207

16 ----- 16

208 TGCTCCAGCTGGCGGACAGCATGCTTCGTGCGATGCCGTATACCCAGGACTGGC 267

17 -----AsnSerSerGluArgGluAspCysAsnAsnGlyGluProPro 30

268 GCACTCACAGCTTCCAGAAATCCACAGAGAGAAAGACTGTAATATGGCGAACCCCT 327

31 ArgLysIleProGluLysAsnSerLeuArgGlnThrTyrAsnSerCysAlaArgLeu 50

328 AGGAAGATAATACCAAGAGAAAGAGCTGTGTGCAAGTATTTAGCAGCTGTGCCAGGCT 387

51 CysLeuAsnGlnGluThrValCysLeuAlaSerThrAlaMetLysThrGluAsnVal 70

388 TGCAATAACCAAGACAGACTATGCTTAAACAGCACTGTCTATCAAGACTGAAATTTGTG 447

71 AlaLysThrLysLeuAlaAsnGlyThrSerSerMetIleValProLysGlnArgLysLeu 90

448 GCCAAGCCAAACTTGGCAATGGCACTTCCAGCATGATGTGCCCAAGCAGCGGAATC 507

91 SerAlaSerTyrGluLysGluLysGluLeuCysValLysTyrPheGluGlnTrpSerGlu 110

508 TCAGCAAGCTATGAGAAAGAAAGAGCTGTGTGCAAGTATTTAGCAGCTGTGTCAGAG 567

111 SerAspGlnValGluPheValGluHisLeulleSerGlnMetCysHisTyrGlnHisGly 130

568 TCTGATCAAGTGGAAATTTGTAGAACACCTTATATCCAAATGTGTCACTACCAAGTGG 627

131 HisIleAsnSerTyrLeuLysProMetLeuGlnArgAspPheIleThrAlaLeuProAla 150

628 CACATCAACTCTACTATAAACTATGCTGCAGAGGGATTTTCATAACTGCATGCACGCA 687

151 ArgGlyLeuAspHisIleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCys 170

688 CGGGGTCTGGACCACATCGCTGAGAACAATTTCTGTATCTTGGACGCCAAGTCACTGTGT 747

171 AlaAlaGluLeuValCysLysGluTrpTyrArgValThrSerAspGlyMetLeuTrpLys 190

748 GCTGCTGAGCTCGTGTGCAAGGAATGGTACCGCGTGACGTGGACGGCATCTGTGSGAAA 807

191 LysLeuIleGluArgMetValArgThrAspSerLeuTrpArgGlyLeuAlaGluArgArg 210

808 AAGCTCATCGAGAGGATGCTCAGGACGGACTCTCTGTGCGGAGGCTGCGAGAGCGCAGA 867

211 GlyTyrGlyGlnTyrLeuPheLysAsnLysProAspGlyAsnAlaProProAsnSer 230

868 GGCTGGGACAGTACTTATTCAAACCAAACTCTCTGATGAGAACGCTCTCCCAACTCC 927

231 PheTyrArgAlaLeuTyrProLysIleIleGlnAspIleGluThrIleGluSerAsnTrp 250

928 TTTTATAGAGCGCTTATCCTAAATCATACAAGACATTGAGACAATAGAGTCCAATTGG 987

251 ArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSerGluThrSerLysGlyVal 270

988 AGATGGGCGGACATAGTTTACAGAGAATCCACTCCGGAGTGAACCAAGTAAAGGGGTT 1047

271 TyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLys 290

1048 TACTGTTTACGTACGACGACCCAGAGATAGTCACGGGCTTCGAGACAACACCATCAAG 1107

291 IleTrpAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerVal 310

1108 ATCTGGGATAAAAGCACACTGGAATGCAAGCGGATTTCTCAGGGGCCACACGGGCTCCGTC 1167

311 LeuCysLeuGlnTyrAspGluArgValIleIleThrGlySerSerAspSerThrValArg 330

1168 CTGTGTTCCAGTACGATGAGAGGGTGATCATCAGAGCTCTCAGACTCCACCGTCAAG 1227

331 ValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisCysGluAlaVal 350

1228 GTGTGGGATGTAATGTCAGGTGAGATGCTTAAACACATTGATTCAACCTGTGAAGCCGTT 1287

351 LeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSerLysAspArgSerIleAla 370

1288 CTGCACCTCGCTTCAATAATGGCATGATGGTGGTGTCTCCAAAGACCGTTTCCATCGCT 1347

371 ValTrpAspMetAlaSerProThrAspIleThrLeuArgArgValLeuValGlyHisArg 390

1348 GTGTGGGATATGGCTTCCCAACTGACATCACCTCAGAGGGGTGCTGGTGGGACACCGA 1407

391 AlaAlaValAsnValValAspPheAspAspLysTyrIleValSerAlaSerGlyAspArg 410

1408 GCTGCGGTCAATGTTGTAGACTTTGATGACAAGTACATCGTTTCTGCGCTCTGGAGATAGA 1467

411 ThrIleLysValTrpAsnThrSerThrCysGluPheValArgThrLeuAsnGlyHisLys 430

1468 ACCATAAAGGTGTGAACAACAAGTACCTGTGTAATTCGTAAGGACCTTAATGGGCAAG 1527

431 ArgGlyIleAlaCysLeuGlnTyrArgAspArgLeuValValSerGlySerSerAspAsn 450

1528 CGTGGCATCGCTGTTTTCAGTACAGACAGCGCTGGTGGTGAGCGGCTCTCTGACAAC 1587

451 ThrIleArgLeuTrpAspIleGluCysGlyAlaCysLeuArgValLeuGluGlyHisGlu 470

1588 ACCATCAGCTGTGGGACATAGAGTGTGGAGCATGCTCGAGTGTGTGAGGGGCCATGAG 1647

471 GluLeuValArgCysIleArgPheAspAsnLysArgIleValSerGlyValaTyrAspGly 490

1648 GAGTGGTACGCTGCATTTCGATTTGATACAAAGAGTAGTGAGCGGAGCCCTATGATGGG 1707

491 LysIleLysValTrpAspLeuValAlaAlaLeuAspProArgAlaProAlaGlyThrLeu 510

1708 AAAATTAAAGTGGGATCTTATGCTGCTTTGGACCGCGTGCTCCAGCGGACTCTC 1767

511 CysLeuArgThrLeuVal 516
 |||||
 1768 TGTCGCGACACTTGTG 1785

AK041532 3824 bp mRNA linear HTC 05-DEC-2002
 Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched library, clone:A630019L11 product:F-BOX/WD-REPEAT PROTEIN 1B (F-BOX AND WD-REPEATS PROTEIN BETA-TRCP2) homolog [Homo sapiens], full insert sequence.

AK041532.1 GI:26334548
 HTC, CAP trapper.

Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)

99279253
 10349636

2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)

20499374
 11042159

3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)

NE 5030913
 ED 11076861

4
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuell, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staib, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamly, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodrigues, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S. and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)

NE 21085660
 ED 11217851

5
 The FANTOM Consortium and the RIKEN Genome Exploration Research

Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)

6 (bases 1 to 3824)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, N., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site for further details.
 URL:http://genome.gsc.riken.go.jp/
 URL:http://fantom.gsc.riken.go.jp/

FEATURES
 Location/Qualifiers

1..3824
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM DB:A630019L11"
 /db_xref="taxon:10090"
 /clone="A630019L11"
 /tissue_type="thymus"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="3 days neonate"
 169..1758
 /protein_id="BAC30975.1"
 /db_xref="GI:26334549"
 /translation="MEPDSVIEDKTIELMNTSVMEDQNEDESPKSKALWQISNGTSSV IVSRKPSSENGYQKRDLCIGYFDQMSDQVEFVEHLISRMCHYOHGHINSYVKPML QRDPTFALPEQGLDHTAENILSYLDARSALCAELCKEWRVISEGMLWKKLERVVR TDFPLKGLSERRGQDYLFPKRPDTPGPNFSYRSYLPKIIQDIETIESNWRGHRHLQ ERVIVSGSDSVYQYDDDKIISGLRDSIKWDKSSLECLKVLGTGSGVLCLOYD ASATDITTLARVLVGHRAAVNVVDFDKYIVSAGDRTIKVSTSTCEPVTNLGHKRG IACLOYEDRLVVGSDNTIRLWDIECGACLEVLGHEELVRCIFREDKRIYSGVDG KIKVWDQLQALDPRAPASTLCRLTVEHSGRVRFLQDFEQIISSHDDTILLWDFLN VPSSAQNETRSPSRITYISR"

polyA_signal
 3805..3810
 /note="putative"

polyA_site
 3824
 /note="putative"

BASE COUNT
 885 a 915 c 982 g 1042 t
 ORIGIN

at Scores:

6.77e-253 Length: 3824
 2395.00 Matches: 449
 Similarity: 86.47% Conservative: 43
 MisMatches: 37
 Indels: 40
 Gaps: 3

01-168B-2 (1-569) x AK041532 (1-3824)

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
 169 ATGGAGCC---GACTCGGTGATTGAAGACAAGACCATCGAGCTCATGAACACATTCAGTT 225
 21 ArgGluAspCysAsnAsnGlyGluProProArgLysIleLeuProGluLysAsnSerLeu 40
 226 ATGGAAGATCAAAATGAAGATGAGTCCCAAGAAA----- 261
 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
 261 ----- 261
 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
 262 -----AGTGTCTTTGGCAGATCAGTAATGGAACGTCA 294
 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
 295 TCTGTGATGTCTCAGAAAGAGCGCGTCAGAGGGAACTACCAAGAAAGAAAGACATTG 354
 101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
 355 TGCAATTAAGTACTTTGACCAAGTGTCTCAATCAGATCAGGTGGAAATTTGTGGACATCTT 414
 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
 415 ATCTCAGGATGTGCATATCAGATGAGCATATTAATCTTACCTGAAGCCCATGTTG 474
 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
 475 CAGCGGGACTTTATCACTGCTTTACAGAGCAGCGCTTAGATCAGATGAGCAAAACATT 534
 161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTyr 180
 535 CTCCTCCTACCTGGATGCCAGGTCTCTGTGTGCGACAGAGCTGGTGTGTAAGAATGGCAG 594
 181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
 595 CGAGTGATCTCAGAGGGATGCTTTGGAGAAAGCTGATTGAGAGGATGGTGGCCAGCCGAC 654
 201 SerLeuTyrArgGlyLeuAlaGluArgArgGlyTyrGlyGlnTyrLeuPheLysAsnLys 220
 655 CCTCTCTGGAAGGAGCTCTCAGAAAGAGAGGCTGGGATCAGTACCTGTTTAAACACAGA 714
 221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleLeu 240
 715 CCTACAGATGGC-----CCTCCCAACTCATTTATAGATCATTTATACCCAAAGATTATC 768
 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
 769 CAGGACATAGACCATAGATCAATCACTGGCGGTGGAGCAGACACAACTTCAGAGGATC 828
 261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
 829 CAGTGGCGCTCTGAAAATAGTAAGGTGTCTACTCTTTGCAATATGATGATGACAAAATT 888
 281 ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300
 889 ATCAGTGGCTCCGGGACAACTCTATCAAGATCTGGGATAAAGCAGCTTGAATGTTG 948
 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320

949 AAAGTGTAAACGGGCCACACAGGCTCTGTCTCTCCAGTATGATCAGCGAGTCATT 1008
 321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyCysMetLeu 340
 1009 GTAACCTGGTCTTTCAGACTCCACGGTGAGAGTCTGGGATGTGAACACACTGGTGGGTCTC 1068
 341 AsnThrLeuIleHisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
 1069 AACACACTCATCCACACCAATGAGCGCTACTGCACTTACGCTTCAGCAATGGACTGATG 1128
 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
 1129 GTGACTTGTTCACAGGACGTTTCCATTCGCGTGTGGGACATGGCTTCTGCCACCGATATC 1188
 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
 1189 ACTTTACGCGGTGTCTGTGTGGCCACCGTGTCTGTCATGTAGTAGACTTTTGATGAT 1248
 401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
 1249 AATACATCGTCTGTCTTCAGGACAGGACCAATTAAGTGTGGAGCAGGACATGT 1308
 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
 1309 GAGTTGTCCGCACTCTGAATGGGCACCAAGGAGGACATCGCTGTCTGCAGTACCGGAC 1368
 441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
 1369 CGGCTTGTGTAGTGGATCATCAGATTAATACCATCCGTTATGGATATTGAATGTGT 1428
 461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
 1429 GCCTGTTTAAAGAGTCTCAGAGGGCAGCAAGAAATTTGGTCCGTGCATCCGTTTGATAAC 1488
 481 LysArgIleValSerGlyValArgGlyLysLysValTrpAspLeuValAlaAla 500
 1489 AAGAGATTGTGATGGCCCTTATGATGGGAAGATTAAAGTCTGGACTTGCAGGCTGT 1548
 501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
 1549 CTTGACCTCGGGCCCCAGCAAGACATTTGTCTGTGCGACCTTGGTGGACACTCTGA 1608
 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerSerHisAspAspThr 540
 1609 CBTGTCTTTCGCTGCGATTTGATGATTTTCAGATTCATCAGCAGCTCCCATGATGACACT 1668
 541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
 1669 AITTTGATTTGGATTCTTAATGTGCTCCAGTGGCCCAATGACACCGCTCTCC 1728
 561 SerArgThrTyrThrTyrIleSerArg 569
 1729 TCCAGAACCTACACCTATATCTCCAGA 1755

RESULT 5

CD500760 1136 bp mRNA linear EST 12-JUN-2003
 CDA46-H04.x1d-t SHGC-CDA Gasterosteus aculeatus cDNA clone
 LOCUS CDA46-H04 5', mRNA sequence.

DEFINITION

CD500760

ACCESSION

CD500760.1 GI:31427791

VERSION

EST.

KEYWORDS

Gasterosteus aculeatus (three spined stickleback)

SOURCE

Gasterosteus aculeatus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;

Gasterosteidae; Gasterosteus.

1 (bases 1 to 1136)

Kingsley,D.M., Peichel,C., Grimwood,J., Dickinson,M.,

Schmutz,J. and Myers,R.M.

Expressed sequence tags from Gasterosteus aculeatus

Unpublished

JOURNAL

Contact: Kingsley, DM
 HHMI and Department of Developmental Biology
 Stanford University School of Medicine
 Beckman Center B300, 279 Campus Drive, Stanford, CA 94305-5329, USA
 Tel: 650 725 5954
 Fax: 650 725 7739
 Email: kingsley@cmmg.stanford.edu
 Plate: 46
 High quality sequence stop: 862.

Location/Qualifiers
 1. .1136
 /organism="Gasterosteus aculeatus"
 /mol_type="mrna"
 /strain="Salinas river, CA"
 /db_xref="taxon:69293"
 /clone="CDA46-H04"
 /sex="mixed male and female"
 /tissue_type="heads and internal organs combined"
 /dev_stage="adult"
 /clone_lib="SHGC-CDA"
 /note="Vector: lambda ZAP Express/pBK-CMV; Site 1: EcoRI (5' adaptor); Site 2: XhoI (3' linker primer). The mixed organ cDNA library was generated using the ZAP-cDNA method by Stratagene. First strand cDNA synthesis was primed with a 50 bp linker primer containing an oligo dT sequence preceded by a synthetic XhoI site. 5 prime adaptors were used containing an EcoRI cohesive end. The finished cDNAs were inserted in to the ZAP express vector unidirectionally in the sense orientation with respect to the lacZ promoter of pBK-CMV. An amplified library was prepared from approximately 3 million primary clones in the lambda ZAP Express vector. In vivo excision was then used to generate individual pBK-CMV phagemid clones for EST sequencing."
 262 a 304 c 302 g 266 t 2 others
 nt Scores:
 1.69e-161 Length: 1136
 1561.00 Matches: 301
 Similarity: 92.98% Conservative: 17
 cal Similarity: 88.01% Mismatches: 23
 Indels: 51.45%
 Gaps: 14

01-168B-2 (1-569) x CD500760 (1-1136)

229 AsnSerPheTyArgAlaLeuTyrProIysIleIleGlnAspIleGluThrIleGluSer 248
 3 AACTCTACTACCACTCCCTTTATCCCAAGATCATCCAAAGACATAGAGACTATTGAGGCA 62
 249 AsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSerGluThrSerLys 268
 63 AATTGCGGTGTGGCAGACACAACTTTCAGCGGATTCAGTGTGCTCAGAAACAGTAA 122
 269 GlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLeuArgAspAsnThr 288
 123 GGGGTTTACTGTCTCCAGTACGATGACGACAAAGATCATCAGTGGGCTTAGGACAAATTC 182
 289 IleLysIleTyrAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGly 308
 183 ATCAAGATCTGGATTAAGCAGTCTCTGGAGTGTCTGAAGATACTGACCGGTCCACAGGGA 242
 309 SerValLeuCysLeuGlnTyrAspGluArgValIleIleThrGlySerSerAspSerThr 328
 243 TCAGTGTGTGTCTCAGATATGCGAGAGGGTCATCGTACCGGCTCTCTGACTCGACT 302
 329 ValArgValTyrAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisCysGlu 348
 303 GTGAGGTTGTGGAGGTCACTACGGGTGAGGTACTGAACACGCTGATCCACCAACAGAG 362
 349 AlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSerLysAspArgSer 368

Db 363 GCGTTTCTTCACTACGCTTGCCTCAACGGATTGATGTCACCTGCTCCAAGGACCGCTCG 422
 Qy 369 IleAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgArgValLeuValGly 388
 Db 423 ATGCGGTCTGGGACATGGCTCCCGACTGATCATGCTAGCCGCTCTCTGCTGGGA 482
 Qy 389 HisArgAlaAlaValAsnValValAspPheAspLysTyrIleValSerAlaSerGly 408
 Db 483 CACCGGGTCTGTCAATGTGTGTCGACTTTGATGACAAATACATGTCGCGCTCAGGG 542
 Qy 409 AspArgThrIleLysValTrpAsnThrSerThrCysGluPheValArgThrLeuAsnGly 428
 Db 543 GACCGTACCAATAAAGGTATGGAGCAACAGTACCTGTGAGTTGTGCGCACATAAATGTT 602
 Qy 429 HisLysArgGlyIleAlaCysLeuGlnTyrArgAspArgLeuValValSerGlySerSer 448
 Db 603 CATAAGAGAGGATCGCTGTCTACAGTACAGAGATCGCTGGTGTGAGGCTCATCG 662
 Qy 449 AspAsnThrIleArgLeuTrpAspIleGluCysGlyAlaCysLeuArgValLeuGluGly 468
 Db 663 GACACACAAATCCGTTATGGGATATAGAGTGTGGGGCGTGTTCGAGTGTCTGGAAGGT 722
 Qy 469 HisGluGluLeuValArgCysIleArgPheAspAsnLysArgIleValSerGlyAlaTyr 488
 Db 723 CACGAGAGTGTGGTGTGCTTCATTCGCTTTCACATATAAAGGATTCAGCGGTGCTAC 782
 Qy 489 AspGlyLysIleLysValTrpAspLeuValAlaAlaLeuAspProArgAlaProAlaGly 508
 Db 783 GAGGCGAGATCAGGTGTGGGACCTGCANGACGCCCTGGA-CCAGAGCCCCCAGCCAGC 841
 Qy 509 ThrLeuCysLeuArgThrLeuValGluHisSerGlyArgValPheArgLeuGlnPheAsp 528
 Db 842 ACATTATGTCTGGCAGCTCTAGTGGAGCACTCTGGGGCGGTGTTCGCTCTCCAGTTTGT 901
 Qy 529 GluPheGlnIleValSerSerHisAspAspThrIleLeuIleTrpAspPheLeuAsn 548
 Db 902 GAGTTTCAGATCATCAGCAGTCTTCAGGAGTACCATCTCTGAICTGGGACTTTCCTGAC 961
 Qy 549 ---AspProAlaAlaGlnAlaGluProProArgSerProSerArgThrTyrThrTyrIle 567
 Db 962 GTCTCGACTAATGGCAGTGGG-AAGGAACGGTCTCCCTCCCGCACTACACTTACATT 1020
 Qy 568 SerArg 569
 1021 TTTAGA 1026
 RESULT 6
 CA985560 924 bp mRNA linear EST 06-JAN-2003
 LOCUS AGENCOURT 11076787 NICHDI_XGC_001 Xenopus laevis cDNA clone
 DEFINITION IMAGE:6860098 5', mRNA sequence.
 CA985560
 VERSION CA985560.1 GI:27518227
 KEYWORDS EST.
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 Xenopodinae; Xenopus.
 1 (bases 1 to 924)
 REFERENCE
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LLAM14473 row: 0 column: 09

High quality sequence start: 16

High quality sequence stop: 717.

Location/Qualifiers

1. 924

/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8355"

/clone="IMAGE:6860098"

/tissue_type="oocytes"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NICHD XGC OOI"

/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI;

Cloned unidirectionally. Primer: Oligo dT. Average insert

size 2.2 kb. Constructed by Life Technologies."

236 a 188 c 264 g 236 t

nt Scores:

10.: 1.16e-144 Length: 924

Similarity: 1407.50 Matches: 278

cal Similarity: 96.22% Conservative: 2

atc: 95.53% Mismatches: 5

46.39% Indels: 6

14 Gaps: 1

01-168B-2 (1-569) x CA985560 (1-924)

246 IleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSerGlu 265

31 ATCGAGTCCAACTGGCGCTGGGAGACACAGCTTACCAAGATTCTCGCGGAGTGAA 90

266 ThrSerLysGlyValTrpCysLeuGlnTrpAspGlnLysLeValSerGlyLeuArg 285

91 ACAAGCAAGGGGTATCTCTCGATGATCAGAGATAGTAAAGTGAGCTCAGA 150

286 AspAsnThrIleLysLeuTrpAspLysAsnThrLeuGluCysLysArgIleLeuThrGly 305

151 GATACACCAATTAGATCTGGGATGAAGATACITTTGGAGTGCAGGAGGCTGATGGGT 210

306 HisThrGlySerValLeuCysLeuGlnTrpAspGluArgValIleLeuThrGlySerSer 325

211 CACACTGGGTGAGTCTCTGCTGCAATATGATGAGAGATATCACTTACTGCTTCA 270

326 AspSerThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIleHis 345

271 GACTCTACCGTCCCGGTGGGAGCTGAACACAGAGAAATGTGAACACGCTGATTCAC 330

346 HisCysGluAlaValLeuHisLeuArgPheAsnGlyMetMetValThrCysSerLys 365

331 CACTGTGAGGTGCTGCTGCACTTGAGCTTTAATATGCGATGATGTCACCTCTCAA 390

366 AspArgSerIleAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgVal 385

331 GATGCTTCATTCGAGTTTGGGAGATGCGCTCTCGACCGCATATCATTACGAAGATT 450

386 LeuValGlyHisArgAlaAlaValAsnValValAspPheAspLysTyrIleValSer 405

451 CTGGTAGGCCACCGAGTGGTATTCATGCTGAGCTTTGATGACAAATATAGTCTCT 510

406 AlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCysGluPheValArgThr 425

511 GCATCTGTGTGTCGAACAATTAAGGTTTGGAAACACCATGATGATGATTTGTGCGACA 570

426 LeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTrpArgAspArgLeuValSer 445

571 TTGAACGCCCAACAGCGTGGTATTCATGCTTACAGATATCGAGATCGGCTTGTGGT 630

446 GlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGlyAlaCysLeuArgVal 465

631 GGTCTTCTGACACACAAATCAGATTGGGATATTAATGCGGTGATGTTTGGCGGTT 690

466 LeuGluGlyHisGluLeuValArgCysIleArgPheAspAsnLysArgIleValSer 485

691 CTGAGGAGCATGAGAACTGTCGCGTGCATCCGCTTTGATACAGAGATAGTCAGT 750

486 GlyAlaTrpAspGlyLysIleLysValTrpAspLeuValAlaAlaLeuAspProArgAla 505

751 GGAGCATATGACGGGAAATTAAGTGTGGGACTTGTGCTGCTTTGGACCCCGTGCC 810

506 -ProAla-GlyThrLeuCysLeuArg-ThrLeu-ValGluHisSerGlyArgValPhe-- 523

811 GCCTGCAAGGACCCCTGTGTCCTCGGACTCTTTTGGAGCATTCGGGCGAGATTTTC 870

524 -ArgLeuGln-PheAspGluPhe 530

871 CGCCTTGACGTTTTCAGGAGTTT 893

RESULT 7

AL898993/c

LOCUS AL898993 XGC-egg Silurana tropicalis cDNA clone TEG9019e21 3', mRNA

DEFINITION sequence.

ACCESSION AL898993

VERSION AL898993.1 GI:22951385

KEYWORDS EST.

SOURCE Silurana tropicalis (western clawed frog)

ORGANISM

REFERENCE

AUTHORS Taylor, R., Ashurst, J.L., Croning, M.D.R., Zorn, A.M. and Rogers, J.

TITLE Sanger Xenopus tropicalis EST project 2002

JOURNAL Unpublished

COMMENT Contact: Taylor R

Sanger Centre

Hanxton, Cambridgeshire, CB10 1SA, UK

Email: trop@sanger.ac.uk

Sanger Xenopus tropicalis EST project 2001

TROPICALIS_SEQUENCE_ID: TEG9019e21.q1k17

Sequencing primer: T7

This sequence is from a Xenopus Gene Collection (XGC) library

constructed by Aaron M. Zorn.

FEATURES

Source

1..895

/organism="Silurana tropicalis"

/mol_type="mRNA"

/db_xref="taxon:8364"

/clone="TEG9019e21"

/dev_stage="egg"

/lab_host="Escherichia coli XL1-blue"

/clone_lib="XGC-egg"

/notes="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA

was oligo dT primed from Sug of poly A+ RNA from egg.

EcoRI-NotI cut cDNA was then ligated into pCS107 with

EcoRI at the 5' end and NotI at the 3' end"

BASE COUNT 232 a 232 c 199 g 232 t

ORIGIN

Alignment Scores:

Pred. No.: 9.46e-143 Length: 895

Score: 1390.00 Matches: 264

Percent Similarity: 98.15% Conservative: 1

Best Local Similarity: 97.78% Mismatches: 5

Query Match: 45.81% Indels: 0

DB: 9 Gaps: 0

US-09-601-168B-2 (1-569) x AL898993 (1-895)

Oy 300 LysArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTrpAspGluArgVal 319

Db 894 AAACGAGTGTGATGGGCCACATGGGTGAGTCTCTGCTGCGATGATGATGAGAGTA 835

Oy 320 IleIleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMet 339

834 ATCAATACCGGCTCTTCTGACTACTACCGTCGGGTTGGGACGTAAACACAGAGAAATG 775
 340 LeuAsnThrLeuIleHisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMet 359
 774 TTGAACACTCTGATTCAACCACTGTGAGGCTGTGTACATTGAGGTTTAAACAATGGTATG 715
 360 MetValThrCysSerLeuAspArgSerIleAlaValTrpAspMetAlaSerProThrAsp 379
 714 ATGGTCACCTGCTCCAAAGATCGTTCCATTGCAGTGTGGACATGGGCTCTGCCACAGAT 655
 380 IleThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAsp 399
 654 ATCAATACGAAGAGTTCTAGTAGCCACCGAGCTGCCGTAATGTGGGACTTTTGAT 595
 400 AspLysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThr 419
 594 GACAAGTACATAGTCTCGCTCTGGTGATCGAACATAAAGGTTTCGAAACACAGTACA 535
 420 CysGluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArg 439
 534 TGTGAATTGTGGCGACATTGAATGGCCACAGCGTGTATTGCTTACAGTATCGA 475
 440 AspArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCys 459
 474 GATAGGCTGTGGTGAAGTGTCTCTGATACCAATCAGATTGTGGGACATTGAATGC 415
 460 GlyAlaCysLeuArgValLeuGluGlyHisGluLeuValArgCysIleArgPheAsp 479
 414 GGAGCATGTTTGGGGTCTCGAGGACATGAAGAACTGGTTTCGATCGATCTTGTAT 355
 480 AsnLysArgIleValSerGlyValTyrAspGlyLysIleLysValTrpAspLeuValAla 499
 354 AACACAGATAGTCAGTGGAGCATATACCGGAAATTAAGTGTGGGACCTTGTGCT 295
 500 AlaLeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSer 519
 294 GCTTTAGACCCCGGACCTCGACCGAGGACCTGTCTCCGGACTCTGTGGAGCATCG 235
 520 GlyArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAsp 539
 234 GGCAGATTTTCCGCTCGCTGATTTGATCAGTTCAGATAGTAACAGCTCTCAGCAGC 175
 540 ThrIleLeuIleTrpAspPheLeuAsnAspProAlaGlnAlaGluProProArgSer 559
 174 ACTATCCCATCTCGGATTTTCTAACACCGCAGTCGACATGCAGATCAACGCGTTC 115
 560 ProSerArgThrTyrThrTyrIleSerArg 569
 114 CTTTCTAGACATACACGTACATCTCCAGA 85

8 2
 CA980902 966 bp mRNA linear EST 27-FEB-2003
 ION AGENCOURT 11279402 Wellcome CRC psk egg xenopus laevis cdna clone
 ON IMAGE:6869574 5', mRNA sequence.

CA980902.1 GI:27513556

S EST.

Xenopus laevis (African clawed frog)

ISM Xenopus laevis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 Xenopodinae; Xenopus.

1 (bases 1 to 966)

RS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

Tumor Gene Index, Cancer Genome Anatomy Project (CGAP),

AL Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgaps-r@mail.nih.gov

Tissue Procurement: N. Garrett, P. LeMaire, A.M. Zorn, and J.B.

Gordon (Wellcome/CRC Institute)

cdna Library Preparation: N. Garrett, P. LeMaire, A.M. Zorn, and

J.B. Gordon (Wellcome/CRC Institute)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: XGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Place: LUM4498 row: j column: 05
 High quality sequence start: 15
 High quality sequence stop: 619.
 Location/Qualifiers

FEATURES

source

1. 966

/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8355"

/clone="IMAGE:6869574"

/tissue_type="egg"

/lab_host="DH10B (phage-resistant)"

/clone_lib="Wellcome CRC psk egg"

/note="Vector: pBluescript SK-; Site 1: NotI; Site 2:

EcoRI; cDNAs were oligo-dT primed and directionally

cloned. Library was constructed by N. Garrett, P. LeMaire,

A.M. Zorn, and J.B. Gordon (Wellcome/CRC Institute).

Note: This is a Xenopus Gene Collection (XGC) library."

BASE COUNT 256 a 202 c 269 g 239 t

ORIGIN

Alignment Scores:

Pred. No.: 1.32e-138 Length: 966

Score: 1353.00 Matches: 275

Percent Similarity: 91.80% Conservative: 5

Best local Similarity: 90.16% Mismatches: 19

Query Match: 44.59% Indels: 8

DB: 14 Gaps: 3

US-09-601-168B-2 (1-569) x CA980902 (1-966)

QY 228 ProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleGluThrIleGlu 247
 Db 10 CCAAAATTCATTCTCCAGAGCGCTTACCCAAAATTTATTCAGACATAGAGCAATCGAG 69
 QY 248 SerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSerGluThrSer 267
 Db 70 TCCAACTGGCGCTCTGGGACACAGCTTACAAGAAATTCACCTCCCGGAGTGAACAAGC 129
 QY 268 LysGlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLeuArgAspAsn 287
 Db 130 AAAGGGGTATACCTCTCGAGTACGATGATCAGAGATAGTAAAGTGGACTCAGAGATAC 189
 QY 288 ThrIleLysIleTrpAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThr 307
 Db 190 ACCATTAAAGATCTGGGATAAGAAATCTTTGGAGTCAAGCGAGTGTGGTGCACACT 249
 QY 308 GlySerValLeuCysLeuGlnTyrAspGluArgValIleIleThrGlySerSerAspSer 327
 Db 250 GGGTCAGTTCTCTCTCTGCATATATGATGAGAGAGTAAATCATTAATGGCTCTTCAGACTCT 309
 QY 328 ThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisHisCys 347
 Db 310 ACCGTCCGGGTGTGGGACGTGAACACAGGAGAAATGTTGAACACAGCTGATTCACCACTGT 369
 QY 348 GluAlaValLeuHisLeuArgPheAsnAspGlyMetMetValThrCysSerLysAspArg 367
 Db 370 GAGGCTGTGGCTGACCTTGAGGTTTAATATGGCATATGATGTCACCTGCTCCCAAGATCGT 429
 QY 368 SerIleAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgValLeuVal 387
 Db 430 TCCATTGCAGTTTGGGACATGGCTCTCGCCGACGATATCATCATACGAGAGTCTGGTA 499
 QY 388 GlyHisArgAlaAlaValAsnValValAspPheAspLysTyrIleValSerAlaSer 407
 Db 490 GGCACCCGAGCTGCGGTAATGTGGTGGACTTTTGTATGACAAAGTATATAGTCTCTGCACT 549
 QY 408 GlyAspArgThrIleLysValTrpAsnThrSerThrCysGluPheValArgThrLeuAsn 427

|||||
550 GGTGATCGAACATTAAGGTTTGGACACCACTGATGATTTGTCGGACATTGAAC 609
|||||
428 GlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAspArgLeuValValSerGlySer 447
|||||
610 GGCCACAAAGCGTGGTATTGCATGCTTACAGTATCAGATCGGCTTGGTGGTGGTCT 669
|||||
448 SerAspAsnThrIleArgLeuTrpAspIleGluCysGlyAlaCysLeuArgValLeuGlu 467
|||||
670 TGTGACAC-ACAAATCAGATTGGGATATTGATGCGGTGATGTTTCCGGGGTCTGGAA 728
|||||
468 GlyHisGluGluLeuValArgCysIleArgPheAspAsnLysArgIleValSerGlyAla 487
|||||
729 GGACATGAAGAACTGGTCGGTGCATCCGCTTTGATAC-AAGAGAATAGTCAGTGGAGCA 787
|||||
488 TyrAsp-GlyLysIleLysValTyrAspLeu-ValAlaAlaLeuAspProArgAla---P 506
|||||
788 TATGACGGGAAAAAATAAGTGTGGGAACCTTTGTCGCTCTTTTGGGCCCCCGGGCGGC 847
|||||
506 roAlaGlyThr-----LeuCysLeuArgThrLeuValGluHis---SerGlyArgValP 523
|||||
848 CTCAGGGGACCCCTGGTCTCCCGGAATCCTTGGTGAACATTTCCGGGCAGAGGTTT 907
|||||
523 heArgLeuGln 526
|||||
908 TCCCCCTGGAA 918

9
26

CE244526 776 bp mRNA linear EST 12-FEB-2003
UI-M-FY0-cdqr-k-22-0-UI_r1 NIH_BMAP_FY0 Mus musculus CDNA clone
IMAGE: 6833135 5', mRNA sequence.

CB244526
EST.

CB244526.1 GI:28366170

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 776)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

Cloning Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1..776

/organism="Mus musculus"

/mol_type="mRNA"

/strain="CS7BL/6"

/db_xref="taxon:10090"

/clone="IMAGE: 6833135"

/tissue_type="whole brain"

/dev_stages="embryo 13.5,14.5,16.5,17.5dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH_BMAP_FY0"

(note="Organ: Brain; Vector: pYX-Asc; Site: 1; Ecor I;

Site: 2; Not I; The library was constructed according

Bonaldi, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGACAG. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

BASE COUNT 201 a 185 c 212 g 175 t 3 others
ORIGIN

Alignment Scores:

Pred. No.: 2e-138 Length: 776
Score: 1350.00 Matches: 252
Percent Similarity: 98.06% Conservative: 1
Best Local Similarity: 97.67% Mismatches: 5
Query Match: 44.50% Indels: 0
DB: 14 Gaps: 0

US-09-601-168B-2 (1-569) x CB244526 (1-776)

QY	152	GlyLeuAspHisIleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAla	171
Db	3	GGTCTGGACCACTCGCTGAGAACATTTCTGTACATTTGGACGCCAAGTCACGTGTGTCT	62
QY	172	AlaGluLeuValCysLysGluTyrArgValThr-SerAspGlyMetLeuTrpLysLys	191
Db	63	GCTGAGCTCGTGTGCAAGGAATGTCACCGGTGAGCTCGGACGGCATGCTGTGGAAAAAG	122
QY	192	LeuIleGluArgMetValArgThrAspSerLeuTrpArgGlyLeuAlaGluArgGly	211
Db	123	CTCATCGAGAGGATGGTCAGACGGACTCTCTGTGGCGAGGCTTGGCAGGCGCAGGCG	182
QY	212	TrpGlyGlnTyrLeuPheLysAsnLysProAspGlyAsnAlaProProAsnSerPhe	231
Db	183	TGGGACAGACTATTATTCAAAAACAACCTCTGTATGAGAACGCTCTCTCCCACTCTTT	242
QY	232	TyrArgAlaLeuTyrProLysIleIleGluAspIleGluThrIleGluSerAsnTrpArg	251
Db	243	TATAGAGCGCTTTATCTCTAAATCATACAAGACATTTGAGACAATAGAGTCCAAATTGGA	302
QY	252	CysGlyArgHisSerLeuGlnArgIleHisCysArgSerGluThrSerLysGlyValTyr	271
Db	303	TGTGGCGACATAGTTTACAGAGATCCACTGCGGAGTGAAACAGATAAGGGGTTTAC	362
QY	272	CysLeuGlnTyrAspAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIle	291
Db	363	TGTTTACAGTACGACGACCAAGATAGTCAGCGGCTTCGAGACAACACCATCAAGATC	422
QY	292	TrpAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeu	311
Db	423	TGGGATAAAGCACACTGGATGCAAGCGGATTTCTACGGGCCACACGGGCTCCGTCCTG	482
QY	312	CysLeuGlnTyrAspGluArgValIleIleThrGlySerSerAspSerThrValArgVal	331
Db	483	TGTTCTGCAGTACGATGAGAGGGTATCATCACAGGCTCTCTCAGACTCCACCGTCAGAGT	542
QY	332	TrpAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisCysGluAlaValLeu	351
Db	543	TGGGATGTAAATGCGAGTGAGATGCTAAACACATTTGATTCACCACCTGTGAAGCCGTTCTG	602
QY	352	HisLeuArgPheAsnAsnGlyMetMetValThrCysSerLysAspArgSerIleAlaVal	371
Db	603	CACCTGGCTTCANTAAATGTCATGATGGTGACCTGTTCANAGACCGCTTCCATCGCTGTG	662
QY	372	TrpAspMetAlaSerProThrAspIleThrLeuArgValLeuValGlyHisArgAla	391
Db	663	TGGGATATGGCTTCCCACTGACATCACCTCANGAGGGTGTGTGTGGACACCGAGCT	722
QY	392	AlaValAsnValValAspPheAspAspLysTyrIleValSerAlaSerGlyAsp	409
Db	723	CGGTCATATGTTGTAGACTTTGATGACAGTATCATGTTTCTGCTCTGGAGAT	776

BM944304 810 bp mRNA linear EST 14-MAR-2002
 UI-M-EHOp-bvq-m-18-0-UI_r1 NIH_EMAP_EHOp Mus musculus cDNA clone
 IMAGE:5695673 5', mRNA sequence.
 BM944304
 BM944304.1 GI:19427889
 EST.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 810)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.N.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

```

Seq primer: pyX-5.
Location/Qualifiers
1. 810
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5635673"
/tissue_type="whole brain"
/dev_stage="embryo 18.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NH Etop EHOp"
/note="Organ: Brain; Vector: pyX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
Oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pyX-Asc vector. The library tail
sequence located between the Not I site and the polyA tail
sequence is CAGCCAGCAG. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institute of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."
189 a 184 c 221 g 216 t

```

```

ut Scores:
      3.56E-138      Length:      810
      1348.00      Matches:      252
      Similarity: 97.78%      Conservative: 12
      al Similarity: 93.33%      Mismatches: 6
      tch: 44.43%      Indels: 0
      12      Gaps: 0

1-168B-2 (1-569) x BM944304 (1-810)

276 AspAspGlnIysIleValSerGlyIeuArgAspAsnThrIleIysIleTirAspLysAsn 295
      ::::::::::::::::::::
1  GATGATGACAAAATTATCAGTGGCCCTCGGGACAACTCTATCAAGATCTGGGATAAAAGC 60
      ::::::::::::::::::::

296 ThrIeuGluCysIysArgIleIeuThrClyHisThrClySerValIeuCysIeuGlnTyr 315
      ::::::::::::::::::::

```

Db	61	AGCTTGGAAATGTTTGAAGATGCTAAACGGGCCCCACAGGCTCTGTCTCTCTGCTCTCAAGTAT	130
QY	316	AspGluArgValIleIleThrGlySerSerAspSerThrValArgValTrpAspValAsn	335
Db	121	GATGAGCGAGTCATTTGTAATCTGTTCTTCAGACTCCACGGTGAGAGTCTGGGATGTGAAC	180
QY	336	ThrGlyGluMetLeuAsnThrIleuIleHisHisCysGluAlaValLeuHisLeuArgPhe	355
Db	181	ACTGGTAGGTGCTCAACACTCATCCACCACAAATGAAGCGGTACTGCACTTACCGTTC	240
QY	356	AsnAsnGlyMetMetValThrCysSerLysAspArgSerIleAlaValTrpAspMetAla	375
Db	241	AGCAATGGACTGATGTGACTTGTTCCAAGGACCGTTCCTCCATTCGGGTGGGACATGGCT	300
QY	376	SerProThrAspIleThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnVal	395
Db	301	TCTGCCACCGATATCACTTTAGCCGCGTGTCTGGTTGGCCACCGTGTCTGTCATGTGA	360
QY	396	ValAspPheAspAspLysVtrIleValSerAlaSerGlyAspArgThrIleLysValTrp	415
Db	361	GTAGACTTTGATGATAAATACATCGTGTCTGCTTCAGGACACAGGACCATTAAGTGTGG	420
QY	416	AsnThrSerThrCysGluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCys	435
Db	421	AGCAGCAGCAGCATGTGAGTTTGTCCGCACTCTGAAATGGGCCAACAGCAGGACATCGCTGT	480
QY	436	LeuGlnTrpArgAspArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrp	455
Db	481	CTGAGTACCGCCAGCCGGCTTGTGTAGTGATCATCAGATATATACATCCGTTATGG	540
QY	456	AspIleGluCysGlyAlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCys	475
Db	541	CATATTGAATGTGGTGCTGTTTAAAGATCTCTAGAGGGCCACGAAGAATTTGGTCCGGTGC	600
QY	476	IleArgPheAspAsnLysArgIleValSerGlyAlaTrpAspGlyLysIleLysValTrp	495
Db	601	ATCCGTTTGTATACAGAGAGATTGTCACTGGCGCTATGATGGGAAGATTAAAGTCTGG	660
QY	496	AspLeuValAlaAlaLeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeu	515
Db	661	GACTTGCAGGCTGCTCTTGACCTCGGGCCCCAGCAAGCACATGTTGTCTGCGCACCTTG	720
QY	516	ValGluHisSerGlyArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSer	535
Db	721	GTGGAACACACTGAGCGTGTGTTTCGGCTCGAGTTTGATGAGTTTCAGATCATCAGCAGC	780
QY	536	SerHisAspAspThrIleLeuIleTrpAsp	545
Db	781	TCCCATGATGACACTATTGTGTTGGGAT	810
RESULT 11			
EQ444188			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1..778

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:5709077"

/tissue_type="whole brain"

/dev_stage="embryo 15.5 dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NH BMAP EXO"

/note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 5:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction. ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTGCGTGGAA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

223 a 181 c 204 g 169 t 1 others

UNT

nt Scores:

O.: 1.96e-136 Length: 778
 1332.00 Matches: 253
 Similarity: 98.07% Conservative: 1
 cal Similarity: 97.68% Mismatches: 5
 atch: 43.90% Indels: 1
 13 Gaps: 0

01-168B-2 (1-569) x BQ444188 (1-778)

63 AlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSerSerMet 82
 1 GCTATGAGACTGAAATGTGTGGCCAAAGCCAAACTTGCCTAATGGCACTTCAGCATG 60
 83 IleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluCysVal 102
 61 ATTTGCCCAAGCAGCGGAACTCTCAGCAAGCTATGAGAGGAAAGGAGCTGTGTGTC 120
 103 LysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeuIleSer 122
 121 AAGTATTTTGACGAGTGTCTGAGTCTGATCAAGTGGATTTGTAGAACCTTATATCC 180
 123 GlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeuGlnArg 142
 181 CAAATGTGTCACTACCAAGTGGGACATCACTCACTTAAACCTATGCTGCAGAGG 240
 143 AspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIleLeuSer 162
 241 GATTTTCAATACTGCATCCAGCAGCGGGTCTGGACACACATCGCTGAGAACATTTGTCA 300
 163 TyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTyrTyrArgVal 192
 301 TACTTGGACGCCAAGTCACTGTGCTCTGAGCTGCTGTGTCAGGAATGTCACCGGTG 360
 183 ThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAspSerLeu 202
 361 ACGTCGGACGGCATCTGTGGAAAAAGCTCATCAGAGGATGGTCAGACGGACTCTCTG 420
 203 TrpArgGlyLeuAlaGluArgGlyTrpGlyGlnTyrLeuPheLysAsnLysProPro 222

Db 421 TGGCGAGGCTGGCAGAGCGCAGAGGCTGGGACAGTACTTATTCAAAACCAACCTCCT 480
 Qy 223 AspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAsp 242
 Db 481 GATGAGAACGCTCTCTCCAACTCTTTTATAGAGCGCTTTATCTTAAATCATACAAGAC 540
 Qy 243 IleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCys 262
 Db 541 ATTGAGACAATAGAGTCCATTTGGAGATGTGGCGGACATAGTTTACAGAGATCCACTGC 600
 Qy 263 ArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSer 282
 Db 601 CGAGTGAACAAGTAAAGGGTTTACTGTTTACAGTACGACGACCCAGAAGATAGTCAGC 660
 Qy 283 GlyLeuArgAspAsnThrIleLysIleTyrAspLysAsnThrLeuGluCysLysArgIle 302
 Db 661 GGCCTTCGAGACAC-ACCATCAAGATCTGGANTAAAGCACACTGGAAATGCAAGCGGATT 719
 Qy 303 LeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIleIle 321
 Db 720 CTCACGGGCCACACAGCGGCTCCGTCCTGTGTCTGCTGACGATCAGAGAGGTGATCATC 776
 RESULT 12
 CA971789 1002 bp mRNA linear EST 27-FEB-2003
 LOCUS AGENCOURT 10768214 Wellcome CRC pSK egg Xenopus laevis cDNA clone
 DEFINITION IMAGE:6323303 5', mRNA sequence.
 ACCESSION CA971789
 VERSION CA971789.1 GI:27504443
 KEYWORDS EST.
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 Xenopodinae; Xenopus.
 REFERENCE 1 (bases 1 to 1002)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: N. Garrett, P. LeMaire, A.M. Zorn, and J.B.
 Gurdon (Wellcome/CRC Institute)
 cDNA Library Preparation: N. Garrett, P. LeMaire, A.M. Zorn, and
 J.B. Gurdon (Wellcome/CRC Institute)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: XGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LAM3766 row: d column: 24
 High quality sequence start: 98
 High quality sequence stop: 710.
 Location/Qualifiers
 1..1002
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="IMAGE:6323303"
 /tissue_type="egg"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="Wellcome CRC pSK egg"
 /note="vector: pBluescript SK-; Site 1: NotI; Site 2:
 EcoRI; cDNAs were oligo-dT primed and directionally
 cloned. Library was constructed by N. Garrett, P. LeMaire,
 A.M. Zorn, and J.B. Gurdon (Wellcome/CRC Institute).
 Note: This is a Xenopus Gene Collection (XGC) library."
 BASE COUNT 296 a 207 c 259 g 239 t 1 others
 ORIGIN
 Alignment Scores:


```

325 AACGAGCGCGTCCACCTGGCTTCTGTAAAGGGCTGATGGTCAACAGGAT 384
|||||
367 ArgSerIleAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgValLeu 386
|||||
385 CGCTCCATTGAGTGTGGGATGGCTCTCCACAGACATAAGCCCTCCCGGGTACTC 444
|||||
387 ValGlyHisArgAlaValAsnValValAspPheAspIleValSerAla 406
|||||
445 GTTGACACAGAGCTGACGCTCAATGGTGTATTTGATGACAGTACATGTCTCTGCC 504
|||||
407 SerGlyAspArgThrIleLysValTrpAsnThrSerThrCysGluPheValArgThrLeu 426
|||||
505 TCTGGGATCGGACTATTAAGGTATGGAGTACAAAGCACCTGCGAGTTTGTTCGACACTA 564
|||||
427 AsnGlyHisLysArgGlyIleAlaCysLeuGlnTrpArgAspArgLeuValSerGly 446
|||||
565 AATGTCATAAGCGAAGAAATTGCTGTCTGAGTACAGACCGACTTGTGTGACGGGC 624
|||||
447 SerSerAspAsnThrIleArgLeuTrpAspIleGluCysGlyAlaCysLeuArgValLeu 466
|||||
625 TCTTCAGACACACCATCAGGCTGTGGATATTGAGTGGGAGCATGTTTGGGGTTTG 684
|||||
467 GluGlyHisGluGluLeuValArgCysIleArgPheAspAsnLysArgIleValSer-Gl 486
|||||
685 GGAGGACACGAGGAGCTTGTTCGCTGTATTGCTTTGACAAACAAACGATCGTGGTGG 744
|||||
486 YAlaTrpAspGlyLysIleLysValTrpAspLeuValAlaAlaLeu-AspProArgAlaP 506
|||||
745 AGCGTATGACGGTAAATTTAAGTATGGAGCTTACAGCTGACATTGATCCACGGCTC 804
|||||
506 RoAlaGlyThrLeuCysLeuArgThrLeuVal-GluHisSerGly--ArgValPheArgL 525
|||||
805 CAGCCAGACACTCTGTCTACGCCACCTAGTGGGAACATTCAAGGCGCGAAGTGTTCAGGG 864
|||||
525 euGln-PheAspGlu-PheGlnIleVal-SerSerSerHisAspAspThrIle--LeuI 543
|||||
865 TTCAGTTTGGAGGTTTCCGAATCATCCAAAGTTCCTATGGATGAACCTTATCCCTCA 924
|||||
543 leTrp 544
|||||
925 TCTGG 929
|||||

14
1 CA327271 850 bp mRNA linear EST 27-NOV-2002
ION UI-M-FY0-ccy-f-17-0-UI.r1 NIH_BMAP_P10 Mus musculus cDNA clone
ON IMAGE: 6826098 5', mRNA sequence.
CA327271
S EST.
S CA327271.1 GI:24545369
Mus musculus (house mouse)
ISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CE 1 (bases 1 to 850)
RS NIH-MGC http://mgi.nci.nih.gov/.
AL National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rmail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5 Location/Qualifiers
1. .850

```

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6826098"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_hosts="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_P10"
/notes="Organ: Brain; Vector: pYX-Asc; Site1: EcoR I;
Site2: Not I; The library was constructed according
Ronald, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCAGACAG. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."
BASE COUNT 221 a 190 c 222 g 212 t 5 others
ORIGIN

Alignment Scores:
Pred. No.: Length: 850
Score: 1310.00 Matches: 247
Percent Similarity: 92.63% Conservative: 17
Best local Similarity: 86.67% Mismatches: 19
Query Match: 43.18% Indels: 3
DB: 14 Gaps: 1

US-09-601-168B-2 (1-569) x CA327271 (1-850)
QY 178 GluTrpTyrArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetVal 197
DB 3 GAATGGCAGCGAGTGATCTCAGAAGGGATGCTTTGGAAGAGCTGATTGAGAGGATGGTG 62
QY 198 ArgThrAspSerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPhe 217
DB 63 CGCACCGACCTCTCTGGAGGAGACTCTCAGAAGAGAGGCTGGGATCAGTACCTGTTT 122
QY 218 LysAsnLysProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrPro 237
DB 123 AAAAACAACACCTACAGATGGC-----CTCCCAACTCATTTTATAGATCAATTATACCA 176
QY 238 LysIleIleGlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeu 257
DB 177 AAGATTATCCAGACATAGAGACCATAGAAATCCAACTGGCGGTGGACGACCAACTTG 236
QY 258 GlnArgIleHisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAsp 277
DB 237 CAGAGGATCCAGTCCCGCTCTGAAATAGTAAGGGTGTCTACTGTTTGAATATGATGATGAT 296
QY 278 GlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeu 297
DB 297 NACAAAATTATCATGGCCCTCCGGGCAACTCATCAAGATCTGGGATGAACGACGCTTG 356
QY 298 GluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGlu 317
DB 357 GAATGTTTGAAGTGCTTAACGGGCCACACAGGCTCTCTCTCTCCCTCCAGTATGATGAG 416
QY 318 ArgValIleIleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGly 337
DB 417 CGAGTCATTTAATGTTCTTTCAGACTCCACGCTGAGAGTCTGGGATGTGAACACTGGT 476
QY 338 GluMetLeuAsnThrIleHisHisCysGluAlaValLeuHisLeuArgPheAsnAsn 357
DB 477 GAGGTGCTCAACACACACTCATCCACCAATGAAGCCGCTACTGCACCTTACGCTTCAGCAAT 536

```

358 GlyMetMetValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerPro 377
 537 GCAGTGTGAGTGTTCACAGACCGTTCATTCCGCTGGGACATCGCTTCTGCC 596
 378 ThrAspIleThrLeuArgValLeuValGlyHisArgAlaAlaValAsnValAsp 397
 597 ACCGATATCATCTTACGCCGGTGTCTGTGGCCACCGTGTCTGTCATGTAGTANAC 656
 398 PheAspAspLysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThr 417
 657 TTGTGATGATNAATACATCGTGTCTGCTTCAGGACAGGACCAATTAAGTGTGGACACG 716
 418 SerThrCysGluPreValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGln 437
 717 AGCATGTGAGTTTGTCCGCACTCTGAATGGGCACAGCGAGCATCGCTGTCTGCAN 776
 438 TyrArgAspArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIle 457
 777 GTACGCGACCGCTGTGTGTGTAGTGATCATCNAGATATACATC-CGGTTATGGGATATT 835
 458 GluCysGlyAlaCys 462
 836 TGATGTGTGCTGT 850

15
 36 CB520936 744 bp mRNA linear EST 28-MAR-2003
 TION UI-M-GH0-cel-p-18-0-UI.r1 NIH_BMAP_GH0 Mus musculus CDNA clone
 ION IMAGE: 6841315 5', mRNA sequence.
 36 CB520936
 36 CB520936.1 GI:29354291
 36 EST.
 36 Mus musculus (house mouse)
 36 Mus musculus
 36 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 36 NIH-MGC <http://mgi.nci.nih.gov/>.
 36 National Institutes of Health, Mammalian Gene Collection (MGC)
 36 Unpublished
 36 Contact: Robert Strausberg, Ph.D.
 36 Email: cgapbs-remail.nih.gov
 36 Tissue Procurement: Dr. Jim Lin, University of Iowa
 36 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 36 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 36 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 36 Clone Distribution: Distribution information can be found at
 36 <http://genome.utowa.edu/distribution/mousefl.html>
 36 This clone was contributed by the Brain Molecular Anatomy Project
 36 (BMAP)
 36 Seq primer: pYX-5.
 36 Location/Qualifiers
 36 1..744
 36 /organism="Mus musculus"
 36 /mol_type="mRNA"
 36 /strain="C57BL/6"
 36 /db_xref="taxon:10090"
 36 /clone="IMAGE: 6841315"
 36 /tissue_type="whole brain"
 36 /dev_stage="1, 5, and 15 days newborn"
 36 /lab_host="DH10B (T1 phage resistant)"
 36 /clone_lib="NIH_BMAP_GH0"
 36 /note="Organ: Brain; Vector: pYX-Asc; Site: 1; EcoR I;
 36 Site 2: Not I; The library was constructed according
 36 Bernaldo, Lennon and Soares, Genome Research, 6:791-806,
 36 1996. Denatured RNA was size fractionated on a 1% agarose
 36 gel. First strand cDNA synthesis was primed with oligo-dT
 36 Primer containing a Not I site. Double strand cDNA was
 36 size selected according to mRNA size fraction, ligated
 36 with EcoR I adaptor, digested with NotI and then cloned
 36 directionally into pYX-Asc vector. The library tag
 36 sequence located between the Not I site and the polyA tail
 36 is CGAACTGAAT. This library was created for the University

Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

BASE COUNT 197 a 184 c 202 g 160 t 1 others
 ORIGIN

Alignment Scores:

Pred. No.: 1.37e-133 Length: 744
 Score: 1306.00 Matches: 243
 Percent Similarity: 98.79% Conservative: 1
 Best Local Similarity: 98.38% Mismatches: 3
 Query Match: 43.05% Indels: 0
 Ds: 14 Gaps: 0

US-09-601-168B-2 (1-569) x CB520936 (1-744)

QY 145 IleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIleLeuSerTyrLeu 164
 DB 2 ATAACTGCACCTGCCAGCACGGGCTCGACCACATCGCTGAGAACATTCTGTCATACTTG 61
 QY 165 AspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyrArgValThrSer 184
 DB 62 GACGCCAAGTCACTGTGTGTGTGAGCTGTGTGCAAGGAATGGTACCGCGTGAAGCTCG 121
 QY 185 AspGlyMetLeuTrpLysLysLeuIleGluAtrMetValArgThrAspSerLeuTrpArg 204
 DB 122 GACGCGATGCTGTGGAAAAAGCTCATCGAGAGATGGTACGAGGACTCTCTGTGGCGA 181
 QY 205 GlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLysProProAspGly 224
 DB 182 GGCCTGGCAGACGCGACAGGCTGGGACAGTACTTATCAAAAACAACCTCCTGATGAG 241
 QY 225 AsnAlaProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleGlu 244
 DB 242 AACCTCTCTCCAACTCCTTTATAGAGCGCTTTATCCTATAAATCATACAAGACATGAG 301
 QY 245 ThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSer 264
 DB 302 ACAATAGATCCAAATGGAGATGTGGCGACATAGTTTACAGAGAAATCCACTGCCGAGT 361
 QY 265 GluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLeu 284
 DB 362 GAAACAAGTAAGGGGTTTACTGTTACAGTACGACGACGACGACGAGATAGTCAGCGCCTT 421
 QY 285 ArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLysArgIleLeuThr 304
 DB 422 CGAGACACACCATCAAGATCTGGGATAAAGCACACTGGGATGCAAGCGGATCTCAGC 481
 QY 305 GlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIleIleThrGlySer 324
 DB 482 GGCACACACGGCTCCGTCTGTGTCTGCTGATGAGAGGGGTGATCATCAGGCTCC 541
 QY 325 SerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIle 344
 DB 542 TCAGACTCCACCGTCAGAGTGGGATGTAATGCGAGGTGAGATGCTAAACACATTGATT 601
 QY 345 HisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSer 364
 DB 602 CACCACCTGTGAAGCGCTTCTGCACCTCGCTTCAATTAATGGCATGATGGTGACCTGTTC 661
 QY 365 LysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgArg 384
 DB 662 AAAGACCGTTCATTCGCTGTGTGGGATGATGGCTTCCCAACTGACATCACCCCTCANGAG 721
 QY 385 ValLeuValGlyHisArgAla 391
 DB 722 GTGCTGGTGGACACCGAGCT 742

Search completed: October 22, 2003, 17:11:13
 Job time : 2690 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

ein - nucleic search, using frame_plus_p2n model

October 22, 2003, 13:58:17 ; Search time 385 Seconds
(without alignments)
3989.559 Million cell updates/sec

US-09-601-168b-2

score: 3034
e: 1 MDPFAVLQKALFNMNSSE.....PAAQABPPSPRTTYTISR 569

table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

d: 2552756 seqs, 1349719017 residues

umber of hits satisfying chosen parameters: 5105512

DB seq length: 0

DB seq length: 2000000000

ccessing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

line parameters:

frame+ p2n.model -DEV=xlpl
2.1/USPTO.spool_p/US09601168/runat.22102003.100545.17767/app_query.fasta_1.711
rneseq 19Jun03 -QMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPL=0
r=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
5 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
XCAL -OUTPWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
309601168 @CGN 1 1 396 @runat.22102003.100545.17767 -NCPU=6 -ICPU=3
> -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOC=100 -LONGLOG
400UT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAEXT=0.5 -FGAPOP=6
7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

ed. No. is the number of results predicted by chance to have a
ore greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3034	100.0	1707	21	AAA73132	Human beta-transdu
2	3034	100.0	2151	21	AAK86501	cDNA encoding a hu
3	3034	100.0	2151	21	AAAS1229	Human beta-TrCP co
4	3034	100.0	2151	21	AAZ93350	Sequence encoding
5	3034	100.0	2151	21	AAZ93710	F-box protein hBe
6	3034	100.0	2151	22	AAC84610	Human ZF11 protein
7	3034	100.0	2151	24	AAAL41041	cDNA of Human F-bo
8	3034	100.0	2151	24	ABSS1009	Human cDNA encodin
9	3034	100.0	2419	21	AAZ29233	Human cell signal
10	3027	99.8	3220	22	AAH90079	Human bone marrow
11	3006	99.1	2285	22	AAK51715	Human polynucleoti
12	2997	98.8	1707	21	AAA73131	Mouse ubiquitin li
13	2997	98.8	2175	21	AAZ93714	F-box protein FWD1
14	2993	98.6	3622	22	AAH89966	Human bone marrow
15	2932.5	98.6	2366	22	AAK51717	Human polynucleoti
16	2989.5	98.5	3003	22	AAK52699	Human polynucleoti
17	2989.5	98.5	3003	22	AAK52700	Human polynucleoti
18	2989.5	98.5	3003	22	AAK52701	Human polynucleoti
19	2892	95.0	2207	22	AAK51716	Human polynucleoti
20	2384.5	78.6	2884	22	AAI61150	Human polynucleoti
21	2384.5	78.6	4230	21	AAA51228	Human E3 ubiquitin
22	2384.5	78.6	4344	22	AAK52260	Human polynucleoti
23	2384.5	78.6	4360	22	AAI59364	Human polynucleoti
24	2089.5	68.9	2387	23	ABL06081	Drosophila melanog
25	1695	55.5	9129	23	ABL06080	Drosophila melanog
26	1163	38.3	657	24	ABK86903	Human beta TrCP (b
27	1155.5	38.1	951	23	ABK43155	cDNA encoding huma
28	935	30.8	1561	22	AAC84601	Beta-TRCP.N/SKP2.C
29	702.5	23.2	594	23	ABV58856	Human prostate exp
30	689	22.7	479	22	ABA44075	Human breast cell
31	689	22.7	479	22	ABA54530	Human foetal liver
32	689	22.7	479	22	ABA24312	Probe #2778 for ge
33	689	22.7	479	22	AAK02821	Human brain expres
34	689	22.7	479	22	AAK28263	Human bone marrow
35	689	22.7	479	22	AAI12828	Probe #2761 for ge
36	689	22.7	479	22	AAI34184	Probe #2870 used t
37	689	22.7	479	22	AAI02746	Probe #2737 used t
38	689	22.7	479	23	ABS27866	Human liver single
39	689	22.7	479	24	ABS02776	Human genome-deriv
40	661	21.8	4135	23	ABL20843	Drosophila melanog
41	661	21.8	4146	23	ABL20841	Drosophila melanog
42	661	21.8	4158	23	ABL26469	Drosophila melanog
43	644.5	21.2	1881	20	AAK99712	Human 6myc-N-sel-1
44	644.5	21.2	1881	22	AAF24136	Construct 6myc-N-s
45	640	21.1	1799	22	AAH14890	Human cDNA sequenc

ALIGNMENTS

RESULT 1
AAA73132
ID AAA73132 standard; cDNA; 1707 BP.
XX
AC AAA73132;
DT 27-NOV-2000 (first entry)
XX
DE Human beta-transducin repeat containing protein (beta-TrCP) cDNA.
XX
KW Ubiquitin ligase SCF complex; F-box protein; ubiquitination; IkappaB;
KW beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FWD1;
KW gene therapy; colon cancer; beta-transducin repeat containing protein;
KW beta-TrCP; ss.
XX
OS Homo sapiens.
XX

2000166542-A.

- JUN-2000.

- DEC-1998; 98JP-0343437.

- DEC-1998; 98JP-0343437..

AGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

I; 2000-485550/43.

PSDB; AAB12813.

box protein of ubiquitin ligase SCF complex which promotes the ubiquitination of IkappaB or beta-catenin - sclosure; Fig 17; 19pp; Japanese.

e present invention describes an F-box motif protein of ubiquitin ligase SCF complex which promotes the ubiquitination of IkappaB or beta-catenin and is constituted by Skp1 protein, Cull1 protein and a repeat motif and has the amino acid sequence of 45 residues (AAB12811) one of two 569 residue sequences (AAB12812, which is mouse ubiquitin ligase FWD1 protein) and (AAB12813, which is human beta-transducin repeat containing protein (beta-TrCP)). The F-box protein can be used for a gene therapy of colon cancer by being recombined to a virus vector. e present sequence encodes the human beta-TrCP protein from the present invention.

quence 1707 BP; 514 A; 349 C; 414 G; 430 T; 0 other;

nt Scores:

O.: 4.02e-303 Length: 1707
3034.00 Matches: 569
Similarity: 100.00% Conservative: 0
cal Similarity: 100.00% Mismatches: 0
at ch: 100.00% Indels: 0
21 Gaps: 0

01-168B-2 (1-569) x AAA73132 (1-1707)

1 MetAspProAlaGluAlaValLeuGlnGluAlaLeuLysPheMetAsnSerSerGlu 20
1 ATGACCCCGCGAGGCGGTGCTGCAAGAGAGAGGCACTCAAGTTTATGAATTCCTCAGAG 60
21 ArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLysAsnSerLeu 40
61 AGAGAACACTGTAATAATGCGGAACCCCTTAGGAAGATAATACCAGAGAAGAAATCACTT 120
41 ArgGlnThrTyraSerSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
121 AGACAGACATACACACGCTGCGCAGACTCTGCTTAAACCAACAAACAGATGTTTAGCA 180
61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
181 AGCACTGCTATGAGACTGAGATTGCTGCGCAACAAACAACTTGCCANTGGCACTTCC 240
81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrrGluLysGluLysGluLeu 100
241 AGTATGATTGTGCCAAGCAACGAAACTCTCAGCAAGCTATGAAAGGAAAGAAAGAACTG 300
101 CysValLysTyrrPheGluGlnThrSerGluSerAspGlnValGluPheValGluHisLeu 120
301 TGTGTCAAAATACTTIGACAGTGTGTGAGTTCAGATCAAGTGGAAATTTGGGAACATCTT 360
121 IleSerGlnMetCysHisTyrrGlnHisGlyHisIleAsnSerTyrrLeuLysProMetLeu 140
361 ATATCCCAATATGCTCATACCAACATGGGCACATAAATCTCGTATCTTAAACCTATGTTG 420
141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
421 CAGAGAGATTTTCATAACTGCTGCTGCCAGCTCGGGGATTTGGATCATATCCTGAGAACAT 480

Qy 161 LeuSerTyrrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyrr 180
Db 481 CTGTCTATACCTGGATGCCAAATCACTATGTGTGCTGAACCTGTGTGCAAGGAAATGGTAC 540
Qy 181 AtgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
Db 541 CGASTGACCTCTGTATGGCATGTGTGGAAAGAGCTTATCGAGAGAAATGTCAGGACAGAT 600
Qy 201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrrLeuPheLysAsnLys 220
Db 601 TCTCTGTGAGAGGCTGGCAGACGAAAGAGGATGGGCACAGTATTTATTCAAAAACAA 660
Qy 221 ProProAspGlyAsnAlaProProAsnSerPheTyrrArgAlaLeuTyrrProLysIleIle 240
Db 661 CTTCTGTGAGGGAATGCTCTCCCAACTCTTTTATAGAGCACTTTATCTTAAATATATA 720
Qy 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
Db 721 CRAAGACATTGAGACATAAGATCTAATTGGAGATGTGGAAGACATAGTTTACAGAGAATT 780
Qy 261 HisCysArgSerGluThrSerLysGlyValTyrrCysLeuGlnTyrrAspAspGlnLysIle 280
Db 781 CACTGCCGAAGTCAACCAAGCAAGAGAGTTTACTGTTTACAGTATGATGATCGAGAAATA 840
Qy 281 ValSerGlyLeuArgAspAsnThrIleLysIleLeuTrpAspLysAsnThrLeuGluCysLys 300
Db 841 GTAAGCGGCTTCGAGACACCAATCAAGATCTGGATATAAAACACATTCGAATGCAAG 900
Qy 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrrAspGluArgValIle 320
Db 901 CGAATTTCTCAGAGCCATACAGTTCAGTCCCTCTCTCCAGTATGATGAGAGAGTGATC 960
Qy 321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
Db 961 ATAACAGATATCGGATTCAGCGTCAGAGTGGGATGTAAATACAGGTGAAATGTA 1020
Qy 341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
Db 1021 AACACGTTGATTCACCATTTGAAGCAGTTCAGCACTTCGCTTCAATGTTGATGAC 1080
Qy 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
Db 1081 GTACCTGTCTCCAAAGATCGTTCCATGTCTGTATGGGATATGGCTCCCCAACACTGACATT 1140
Qy 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
Db 1141 ACCCTCCGAGGGTGTGCTGCGACACCGAGCTGCTGTCATGTTGTAGACTTTGATGAC 1200
Qy 401 LysTyrrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
Db 1201 AAGTACATGTTTCTGCATCTGGGATAGAACTATTAAGGTATGGAACACAAAGTACTTGT 1260
Qy 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrrArgAsp 440
Db 1261 GAATTTGTAAGGACCTTAAATGGACAAACAGGAGCATTCCTGTTTGCAGTACAGGGAC 1320
Qy 441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
Db 1321 AGSCTGATGTAGTGGCTCATCTGCAACACATCATCAGATTATGGACATAGAAATGTGT 1380
Qy 461 AlaCysLeuArgValLeuGluGlyHisGluLeuValArgCysIleArgPheAspAsn 480
Db 1381 GCATGTTACAGCTGTAGAGGCCATGAGGAATTTGGTCGTTGTATTCGATTTGNTAAC 1440
Qy 481 LysArgIleValSerGlyAlaTyrrAspGlyLysIleLysValTrpAspLeuValAlaAla 500
Db 1441 AAGAGGATAGTCACTGGGCGCTATGATGAAAAATTAAGTGTGGATCTTGTGGCTGCT 1500
Qy 501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
Db 1501 TTGGACCCCGTCTCTCTCGAGGAGCACTCTGTCTACGAGCCCTTGTGGAGCATTCGGA 1560

521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspThr 540
 1561 AGAGTTCGATCAGTTGAGAAATCCAGATTGTCAGTAGTTTCATGATGACACA 1620
 541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
 1621 ATCTCATCTGGGACTTCCTAAATGATCCAGCTGCCAAGCTGAACCCCGCTCCCT 1680
 561 SerArgThrTrpThrIleSerArg 569
 1681 TCTCGAACATACACCTACATCTCCAGA 1707

2

1 X86501 standard; cDNA; 2151 BP.

X86501;

-MAR-2003 (updated)

-SEP-1999 (first entry)

NA encoding a human beta-transducin repeat containing protein.

ta-transducin repeat containing protein; beta-Trcp; Skplp;
 osteosome degradation pathway; Vpu protein; beta-catenin;
 man immune deficiency virus-1; HIV-1; cellular protein; IkappaB;
 ubiquitylation; phosphorylated protein; tumour; apoptosis; Alzheimer's;
 tiviral; antitumour; cell cycle regulation; protein degradation;
 d anti-inflammatory; osteo-articular inflammation; acute inflammation;
 mour necrosis factor; ss.

to sapiens.

Y Location/Qualifiers

70..179

/*tag= a

/product= beta-Trcp

/note= "beta-transducin repeat containing protein"

938969-A1.

-AUG-1999.

JAN-1999; 99WO-FR00196.

-DEC-1998; 98FR-0015545.

-JAN-1998; 98FR-0001100.

ISP) INST PASTEUR.

IRM) INST NAT SANTE & RECH MEDICALE.

szana Seisdedos F, Benarous R, Concordet J, Durand H;

all M, Margottin F;

; 1999-469329/39.

SDB; AAY24054.

, human beta-transducin repeat containing protein and its
 gments useful as, or to screen for, antiviral, antitumour,
 i-inflammatory and anti-Alzheimer's agents

im 7; Page 57-60; 71pp; French.

present sequence encodes a human beta-transducin repeat containing
 tein (beta-Trcp). The protein directs proteins to the proteosome
 radiation pathways. The protein is able to interact with the Vpu
 tein of human immune deficiency virus-1 (HIV-1), cellular proteins
 ppaB or beta-catenin (bc) and/or protein Skplp. The protein controls
 quitylation of phosphorylated proteins and thus their targeting to
 teosomes for degradation. Depending on whether the process is
 ibited or promoted, the result may be delayed breakdown of CD4 (in
 es of HIV-1 infection); increased activity of IkB (and thus reduced
 ivity of NFkappaB) and increased degradation of mutant bc in tumour

CC cells, or increased bc survival (and reduced apoptosis) in Alzheimer's
 CC patients. The beta-Trcp protein, and its active peptide fragments, or its
 CC nucleic acid, are used to screen for anti HIV-1 agents (antivirals),
 CC antitumour agents that disrupt cell cycle regulation or protein
 CC degradation in human tumour cells, and anti-inflammatory agents that
 CC disrupt activation by NFkappaB. Fragments of the protein are also
 CC useful for treating osteo-articular inflammation or acute inflammation
 CC associated with release of tumour necrosis factor.
 CC (Updated on 20-MAR-2003 to correct PA field.)
 XX

SQ Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;

Alignment Scores:

Pred. No.: 5,66e-303 Length: 2151
 Score: 3034.00 Matches: 569
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 20 Gaps: 0

US-09-601-168B-2 (1-569) x AAX86501 (1-2151)

Qy 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
 Db 70 ATGGACCGCGCGAGCGGTGCTGCAAGAGAAGGCACTCAAGTTATGATTCCTCAGAG 129
 Qy 21 ArgGluAspCysAsnAsnGluProProArgLysIleIleProGluLysAsnSerLeu 40
 Db 130 AGAAGAAGCTGTAATATGCGAACCCCTAGGAAGATAATACCAGAGAAGAAATTCACCT 189
 Qy 41 ArgGlnThrTrpAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
 Db 190 AGACAGACATACACAGCTGTGCCAGACTCTGCTTAAACCAAGAACACAGTATGTTAGCA 249
 Qy 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
 Db 250 AGCACTGCTATGAAGACTGAGAATTGTGTGGCCMAAACAAACTTGCCAATGGCACTTC 309
 Qy 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyTGluLysGluLysGluLeu 100
 Db 310 AGTATGATTGTGCCCAAGCAACGGAACCTCTCAGCAAGCTATGAAAGGAAAGGAAGT 369
 Qy 101 CysValLysTyPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
 Db 370 TGTGTCAATACTTTGAGCAGTGGTCAGAGTCAGATCAAGTCGAATTTGTGGAACATCTT 429
 Qy 121 IleSerGlnMetCysHisTyGlnHisGlyHisIleAsnSerTyLysLeuLysProMetLeu 140
 Db 430 ATATCCCAATGTGTCAATACCAACATGGGCACATAAACTCGTATCTTAAACCTATGTTG 489
 Qy 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
 Db 490 CAGAGAGATTTCATACTGCTCTGCCAGCTCGGGGATTGGATCATATCTGCTGAGAACAT 549
 Qy 161 LeuSerTyLysLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 180
 Db 550 CTGTCTATACCTGGATGCCAAATCACTATGTGCTGTGAACTTGTGTGCAAGGAATGGTAC 609
 Qy 181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
 Db 610 CGAGTGACCTCTGATGGCATGCTGTGGAAGAAAGCTTATCGAGAGAATGGTCAGGACAGAT 669
 Qy 201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTTPGlyGlnTyLysLeuPheLysAsnLys 220
 Db 670 TCTCTGTGAGAGGCTGCGACAGACGAGAGATGGGACAGTATTTATTCAAAAACAAA 729
 Qy 221 ProProAspGlyAsnAlaProProAsnSerPheTyArgAlaLeuTyProLysIleIle 240
 Db 730 CCTCTGTGAGGGAATGCTCCTCCCAACTCTTTTATAGAGCACTTATCTCTAAATATA 789
 Qy 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
 Db 790 CAAGACATTTAGACAATAGATCTTAATTTGGAGATGTGGAAGACATAGTTTACAGAGAATT 849

261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspSerGlnLysIle 280
 |||||
 850 CACTGCCGAAGTGAACCAAGCAAGAGGAGTTACTGTGTTACAGTATGATGATCAGAAATA 909
 |||||
 281 ValSerGlyLeuArgAspAsnThrIleLysIleTyrAspLysAsnThrLeuGluCysLys 300
 |||||
 910 GTAAGCGGCTTCGAGACACACACATCAAGATCGGATGAACACACATTCGATGCAG 969
 |||||
 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
 |||||
 970 CGAATTCTCACAGGCCATACAGGTTTCAGTCTCTGCTCCAGTATGATGAGAGATGATC 1029
 |||||
 321 IleThrGlySerSerAspSerThrValArgValTyrAspValAsnThrGlyGluMetLeu 340
 |||||
 1030 ATACAGGATCATCGGATTCACGGTCAGAGTGGGATGTAATACAGGTGAATGCTA 1089
 |||||
 341 AsnThrLeuIleHisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
 |||||
 1090 AACACGTTGATTCACCATTTGTGAAGCAGTTCTGCACCTGGCTTCAATTAATGGCATGATG 1149
 |||||
 361 ValThrCysSerLysAspArgSerIleAlaValTyrAspMetAlaSerProThrAspIle 380
 |||||
 1150 GTGACCTGCTCCAAAGATCGTTCATTTGCTGTGTGGATATGGCTTCCCACTGACAT 1209
 |||||
 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
 |||||
 1210 ACCCTCGGAGGTTGCTGTCGACACCGAGCTGCTCAATGTTGTAGAC-TTGTATGAC 1269
 |||||
 401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTyrAsnThrSerThrCys 420
 |||||
 1270 AAGTACATGTTCTGATCTGCGGATAGAACATATAAGGTATGGAAACACAGTACTTGT 1329
 |||||
 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
 |||||
 1330 GAATTGTGAAGCACTTAATGGACACAACAGGAGCATTCCTGTTTGCAGTACAGGGAC 1389
 |||||
 441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTyrAspIleGluCysGly 460
 |||||
 1390 AGGCTGTGTAGTGTGCTCATCTGCAACACTATCATGATATATGGACATAGAAATGTGT 1449
 |||||
 461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
 |||||
 1450 GCATGTTTACAGTGTAGAGGCCATGAGGAATGCTGCTGTGTTATTCGATTGATAC 1509
 |||||
 481 LysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTyrAspLeuValAlaAla 500
 |||||
 1510 AAGAGGATAGTCACTGGGGCTATGATGGAAAAATTAAGTGTGGGATCTTGTGGCTGCT 1569
 |||||
 501 LeuAspProAspAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
 |||||
 1570 TTGACCCCTGCTCTCGAGGGACACTCTGTCTACGGACCTTGTGGAGCATTCGGA 1629
 |||||
 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
 |||||
 1630 AGAGTTTTCACATACAGTTTGTGAATTCAGATTTCAGATTTCACATGATGACACA 1689
 |||||
 541 IleLeuIleTyrAspPheLeuAsnAspProAlaAlaGlnAlaGluProArgSerPro 560
 |||||
 1690 ATCTCATCTGGGACTTCTTAATGATCCAGCTGCCCAAGCTGAACCCCTTCCCT 1749
 |||||
 561 SerArgThrTyrThrTyrIleSerArg 569
 |||||
 1750 TCTGACATACACCTACATCTCCAGA 1776
 }
 }
 451229 standard; DNA; 2151 BP.
 451229;
 SEP-2000 (first entry)

Human beta-TrCP coding sequence.

E3 ubiquitin ligase; beta-TrCP; F-box; WD protein; I-kappa-B; inhibitor; nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral; anti-inflammatory; immunosuppressive; cytostatic; ds.

Homo sapiens.

Key	Location/Qualifiers
CDS	70..1779
FT	/*tag= a
FT	/product= Human_TrCP

WO200034447-A2.

15-JUN-2000.

10-DEC-1999; 99WO-US29371.

10-DEC-1998; 98US-021060.

(SIGN-) SIGNAL PHARM INC.

(YISS) YISSUM RES & DEV CO.

Manning AM, Mercurio F, Amit S, Ben-eriah Y, Davis M, Hatzubai A; Lavon I, Yaron A;

WPI; 2000-431294/37.

P-PSDB; AAY96697.

Polypeptide enhancing phosphorylated IkapabB ubiquitination useful for treating disorder associated with NF-kappaB activation e.g. cancer, comprising amino acid sequence of human E3 ubiquitin ligase or its variant

Disclosure; Page 72; 77pp; English.

This DNA encodes human beta-TrCP, an F-box/WD protein family member, which has been shown to have homology to human E3 ubiquitin ligase (E3). E3 enhances ubiquitination of phosphorylated I-kappa-B, an inhibitor protein of nuclear factor kappa-B (NF-kappa-B). Understanding I-kappa-B degradation via the ubiquitin pathway is useful for identifying modulators of this process for use in treating diseases associated with activation of NF-kappa-B. In vitro analysis suggests that deletion of the F-box results in a protein that functions as a dominant negative molecule in vivo. Transient over-expression of delta-beta-TrCP (a deletion mutant) inhibited the degradation of endogenous I-kappa-B-alpha in stimulated Jurkat cells, resulting in accumulation of phosphorylated I-kappa-B-alpha. E3 can be used to screen for modulators of NF-kappa-B activity. E3 and beta-TrCP can be used to modulate NF-kappa-B to treat inflammatory diseases, autoimmune diseases, cancer and viral infections.

Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other:

Alignment Scores:

Pred. No.:	5.66e-303	Length:	2151
Score:	3034.00	Matches:	569
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	21	Gaps:	0

US-09-601-168B-2 (1-569) x AAA51229 (1-2151)

Qy 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
 |||||
 Db 70 ATGACCCCGCCGAGCGGTGCTGCAAGAGAGAGGCACTCAAGTTATGAATTCCTCAGAG 129
 |||||
 Qy 21 ArgGluAspCysAsnAsnGlyGluProArgLysIleIleProGluLysAsnSerLeu 40
 |||||
 Db 130 AGAAGACATGTATTAATGGCAACCCCTTAGAGATATACAGAGAGATTCACCT 189
 |||||
 Qy 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
 |||||

190	AGACAGACATACAAACAGCTGTCCAGACTCTGCTCTAAACCAAGAAACAGATATGTTAGCA	249	1270	AAGTACATGTTTCTGCATCTGGGATAGAACTATAAGGTATGGAAACAAGTACTTGT	1321
61	SerThrAlaMetLysThrGluAsnCysValAlaLysThrIleValSerGlyThrSer	80	421	GluPheValArgThrIleuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp	440
250	AGCACTGCTATGAGACTCAGAAATGTGTGCGCAAAACAAACATTCGCCAATGGCACTCC	309	1330	GAATTTGTTAAGGACCTTAAATCGACACAAACAGGACATTCCTGCTTTGCAGTACAGGAC	1389
81	SerMetIleValProLysGlnArgLysIleuSerAlaSerTyrGluLysGluLeu	100	441	ArgLeuValValSerCysLysSerSerAspAsnThrIleArgLeuTyrAspIleGluCysGly	460
310	AGTATGATTTGCCCAAGCAACGGAACCTCTCAGCAAGCTATGAAAGGAAAGGAACTG	369	1390	AGCTGGTGTAGTGGTGGCTCATCTGACCAACACTATCAGATTTATGGGACATAGAAATGTGGT	1449
101	CysValLysTyrPheGluGlnTyrSerGluSerAspGlnValGluPheValGluHisLeu	120	461	AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleLeuArgPheAspAsn	480
370	TGTGTCAATACTTTGAGCAGTGGTCAGAGTCAGATCAAGTGGAAATTTGTGGAACATCT	429	1450	GCAATGTTACGAGTGTGTAGAGGCCATGAGGAATTTGGTGGCTTCTGATTCGATTTGATAAC	1509
121	IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu	140	481	LysArgIleValSerCysLysValAlaTyrAspGlyLysIleLysValTyrAspLeuValAla	500
430	ATATCCAAATGTGTCATACCAACATGGGACATAAACTCGTATCTTAAACCTATGTTG	489	1510	AAAGAGATAGTCAGTGGGGCTTATGATGAAATAATTAAGGTGGGATCTTTGGGCTGCT	1569
141	GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle	160	501	LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly	520
490	CAGAGAGATTTCTAATACTCTCTGCCACTCGGGATTTGGATCATATCGTGTGAGAACAT	549	1570	TTGGACCCCGTCTCTCGAGGACACTCTGTCTACGGACCTTGTGGAGCATTCGCGA	1629
161	LeuSerTyrIleuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTyrTyr	180	521	ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerSerHisAspAspThr	540
550	CTGTCTATCTGGATGCCAAATCACTATGCTGTGCTGAACTTTGTGTGCAAGGAATGGTAC	609	1630	AGAGTTCCTGCACTACAGTTTCAATTCAGATTCAGATTCACATGATGACACA	1689
181	ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp	200	541	IleLeuIleTyrAspPheLeuAsnAspProAlaAlaGlnAlaGluProCysArgSerPro	560
610	CGAGTGACCTCTGATGGCATGCTGTGGAGAAAGCTTATCGAGAGAAATGGTCAGACAGAT	669	1690	ATCTCTATCTGGGACTTCTTAATGATTCAGCTGCCCAAGCTGAACCCCGCTTCCCT	1749
201	SerLeuTrpArgGlyLeuAlaGluArgGlyTyrGlyGlnTyrLeuPheLysAsnLys	220	561	SerArgThrTyrThrTyrIleSerArg	569
670	TCTCTGTGAGAGGCTGCGACAAACGAGAGATGGGACAGTATTTATTCAAAAACAAA	729	1750	TCTCGAACATACACCTTACATCTCCAGA	1776
221	ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle	240	RESULT 4		
730	CTCTCTCGCGGGAATGCTCTCCCAACTCTTTTATAGAGCACTTTATCTCAAAATATATA	789	AAZ93350		
241	GlnAspIleGluThrIleGluSerAsnTyrArgCysGlyArgHisSerLeuGlnArgIle	260	ID	AAZ93350 standard; cDNA; 2151 BP.	
790	CAAGACATTTGAGACATAGAAATCTAAATGGAGATGGAGACATAGTTTACAGAGAAAT	849	AC	AAZ93350;	
261	HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle	280	XX	16-AUG-2000 (first entry)	
850	CAGTCCGGAAGTGAAACAGCAAGAGGTTTACTCTTTACAGTATGATGATCAGAAATA	909	DE	Sequence encoding F-box protein FBP-1.	
281	ValSerGlyLeuArgAspAsnThrIleLysIleTyrAspLysAsnThrLeuGluCysLys	300	XX	F-box protein; FBP; diagnosis; treatment; screening; agonist;	
910	GTAAAGCGGCTTCGAGACACACATCAAGATCTGGGATATAAACACATTTGGAAATGCAAG	969	KW	antagonist; proliferative disorder; differentiative disorder;	
301	ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle	320	KW	breast cancer; prostate cancer; ovarian cancer; cancer;	
970	CGAATTCACAGGCCATACAGGTTTCAGTCTCTGCTCCAGTATGATGAGAGATGATC	1029	KW	small cell lung carcinoma; immune disorder; cardiovascular disorder;	
321	IleThrClySerSerAspSerThrValArgValTyrAspValAsnThrGlyGluMetLeu	340	XX	inflammatory disorder; human; ss.	
1030	ATAACAGGATCATCGGATTTCCAGGTCAGAGTGGGATGTAAATACAGGTTGAAATGCTA	1089	OS	Homo sapiens.	
341	AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet	360	XX	Key	
1090	AACACGTTGATTCACCATTTGTAGACAGTTCTGCATCTGCGTTTCATATGCGCATGATG	1149	PH	Location/Qualifiers	
361	ValThrCysSerLysAspArgSerIleAlaValTyrAspMetAlaSerProThrAspIle	380	FT	70..1779	
11150	GTACCTGTCTCCAAAGATCGTTCCATGTGTGTGGGATATGGCTCTCCCAACTGACAT	1209	ET	/*tag= a	
381	ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp	400	XX	/product= FBP-1	
1210	ACCTCTCGGAGGGTGTGGTGGACACCGAGTGTGTCAATGTTGTAGACTTTGTATGAC	1269	PN	WO200012679-A1.	
401	LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTyrAsnThrSerThrCys	420	XX	09-MAR-2000.	
			XX	27-AUG-1999; 99WO-US19560.	
			XX	28-AUG-1998; 98US-0098355.	
			XX	03-FEB-1999; 99US-0118568.	
			XX	15-MAR-1999; 99US-0124449.	
			XX	(UUNY) UNIV NEW YORK STATE.	
			XX	Chiour DS, Pagano M, Latres E;	
			XX	WPI; 2000-256635/22.	
			XX	P-PSDB; AAY83041.	

ovel nucleic acid for screening compounds useful for treating proliferative and differentiative disorders such as cancer and immune disorders comprises sequences encoding ubiquitin ligases - disclosure; Figure 3b, 3c, 3d; 245pp; English.

leic acids encoding substrate-targeting subunits of ubiquitin ligases with F-box motifs (F-box proteins) are useful for diagnosis of proliferative and differentiated related disorders by measuring 3P gene expression. Cells expressing such proteins or their fragments are useful for screening compounds. The compounds are agonists or antagonists, which are useful for treating a proliferative or differentiative disorder in a mammal such as breast, ovarian and prostate cancer and small cell lung carcinoma and also major opportunistic infections, immune disorders, cardiovascular diseases and inflammatory disorders. FBP protein, analogs, derivatives and their subsequences, anti-FBP antibodies are also useful in diagnosis of the disorders.

quence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;

nt Scores:

O.: 5.66e-303 Length: 2151
 Similarity: 3034.00 Matches: 569
 cal Similarity: 100.00% Conservative: 0
 Mismatches: 0
 Indels: 0
 Gaps: 0

01-168B-2 (1-569) x AAZ93350 (1-2151)

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
 70 ATGACCGCGCGAGGCGGTGCTGCAAGAGAGAGGCACTCAAGTTTATGATTTCTCAGAG 129
 21 ArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLysAsnSerLeu 40
 130 AGAGAAGACTGTATAATAGCGAACCCTTAGGAAGATAATACAGAGAGGAAATTCACATT 189
 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
 190 AGACAGACATACACAGCTGTGCCAGACTCTGCTTAACCAAGAACAGATGATTTAGCA 249
 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
 250 AGCCTGCTATGAAGACTGAGAAATGTGTGCCAAAAACAAACTTGCCAAATGSCACTTCC 309
 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
 310 AGTATGATTGTGCCAAGCAACGGAACCTCTCAGCAAGCTATGAAAGGAAGGAACCTG 369
 101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
 370 TGTGTCAAAATCTTTAGCAGCTGTCAGAGTCAGATCAAGTGGAATTTGTGGAACATCTT 429
 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
 430 ATATCCCAATGTGTCTATACCAATGTGGCAGCAATAAATCGATCTTAAACCTATGTGTG 489
 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
 490 CAGAGAGATTTTCATAACTGCTCTGCCAGCTCGGGATTGGATCATATCGCTGAGACATT 549
 161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 180
 550 CTGTGCATACCTGGATGCCAAATCACTATGTCTGCTCTGAACCTTGTGTGCAAGGAATGGTAC 609
 181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
 610 CGAGTGACCTCTGATGGCATGCTGTGGAAGAGCTTATCAGAGAAATGGTCAGGACAGAT 669
 201 SerLeuTrpArgGlyLeuAlaGluArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220

670 TCTCTGTGGAGAGCTTGGCAGAACGAGAGGATGGGACAGTATTTATTCAAAACAAA 729
 221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
 730 CCTCTGTGACGGGAATGCTCTCCCAACTCTTTTATTAGAGCACTTTATCTAAATTTATA 789
 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
 790 CAAGACATTGAGACAATAGAAATCTAAATGGAGATGGGAAGACATAGTTTACAGAGAATT 849
 261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
 850 CACTGCCAAGTGAACAGCAAGAGAGGTTTACTGTTTACAGTATGATGATCAGAAATA 909
 281 ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300
 910 GTAAGCGGCTTCCAGACAACACAAATCAAGATCTGGGATAAAACACACATTGGAAATGCAAG 969
 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
 970 CGAATTTCTCAGGCCCATACAGGTTCAAGTCTCTGTTCCAGTATGATGAGAGATGATC 1029
 321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
 1030 ATAACAGGATCATCGGATTCACCGTCCAGATGTGGATGTAAATACAGGTGAAATGCTA 1089
 341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
 1090 AACACGTTGATTCAACATTGTGAAGCAGTTCTGCACCTTCGCTTCAATTAATGCGCATGATG 1149
 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
 1150 GTGACCTGCTCCAAAGATCGTTCCATTGCTGTATGGGATATGCGCTCCCAACTGACATT 1209
 381 ThrLeuArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
 1210 ACCCTCGGAGGTGTGTCGGACACCGAGCTGCTGTCATGTTGTGTAGACTTTGATGAC 1269
 401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
 1270 AAGTACATTGTTCTGCATCTGGGATAGAACCTATAAAGGTATGGAAACACAGTACTTCT 1329
 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
 1330 GAATTTCTAAGGACCTTAAATGGACACAAACGAGCAATGCGCTGTTTGCAGTACAGGGAC 1389
 441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
 1390 AGGCTGTAGTAGTGCTCTCTGACACACTATCAGATTATGGACATAGATGTGT 1449
 461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
 1450 GCATGTTTACAGTGTAGAAAGCCCATGAGGAATTTGGTGTGTTGTTGATTCGATTGATAAC 1509
 481 LysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAla 500
 1510 AAGAGGATGTCTAGTGGGCTATGATGAGAAATTTAAAGTGTGGATCTTTGTGCTGTCT 1569
 501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
 1570 TTGACCCCGTGTCTCTCGCAGGACACTCTGTCTACGAGCCCTTGTGGAGCATTCGCGA 1629
 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
 1630 AGAGTTTTTTCGACTACAGTTTGTATGAATTCAGATTGTCCAGTAGTTTCATGATGACACA 1689
 541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
 1690 ATCTCATCTGGGACTTCTTAATGATCCAGCTGCCAGCTGAACCCCTCCCTTCCCT 1749
 561 SerArgThrTyrThrTyrIleSerArg 569

1750 TCTCGACATACACCTACATCTCCAGA 1776

5

0

1Z93710 standard; DNA; 2151 BP.

1Z93710;

-AUG-2000 (first entry)

-box protein hBetaTrCp coding sequence.

iquitin ligase; SCF; F-box protein; targeted degradation;
 stabilisation; proteolysis; drug discovery; gene therapy; cancer;
 icoproten; Huntington's disease; gene knockout; delivery systems;
 man; ss.

no sapiens.

Y Location/Qualifiers

S 70..1779

/*tag= a

/product= F-box protein hBetaTrCp

200022110-A2.

-APR-2000.

-OCT-1999; 99WO-US23705.

FOOT-1398; 98US-0103787.

ARD; HARVARD COLLEGE.

ou P, Howley P;

I; 2000-317970/27.

PSDB; AAY83250.

rgeting degradation of polypeptide useful for treating cancer and
 her proliferative disorders, involves conjugating polypeptide with
 iquitin protein ligase or inhibiting ubiquitination using organic
 upound

aim 10; Page 171; 185pp; English.

e F-box proteins are a family of ubiquitin ligases (SCF ubiquitin
 gasses) which can be used for the targetted degradation of a target
 lypeptide in vivo. Targetted degradation is achieved by expressing
 e ubiquitin ligase in a cell linked to the interaction domain of
 e target polypeptide and thereby recruiting the target polypeptide
 the ubiquitin ligase. Such methods are useful for decreasing or
 reasing the level of a target polypeptide and for creating and
 reasing a destabilized polypeptide which is subjected to SCF
 iated proteolysis. Degrading any desired protein in a cell is
 fful for preventing or treating diseases caused by the presence of
 normal amount of the specific polypeptides, for drug discovery and
 r gene therapy. Diseases treated include cancer, by degradation of
 roproteins, Huntington's disease, other proliferative disorders and
 robial infections. The method provides a quick and easy
 ernative to gene knockout technology. The target polypeptide can
 degraded at all stages, or a specific stage, of development in the
 .ure animal.

quence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;

ut Scores:

Length:	2151
Matches:	569
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	21

US-09-601-168B-2 (1-569) x AA293710 (1-2151)

QY 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
 DB 70 ATGACCCGCCGAGCGGGTCTGCAAGAGAGAGGCACTCAAGTTTATGATTCCTCAGAG 129
 QY 21 ArgGluAspCysAsnAsnGlyGluProArgLysLysLeileProGluLysAsnSerLeu 40
 DB 130 AGAGAGAGACTGTAATAATGGCAACCCCTTAGAGAGATAATACCAGAGAGAAATTCACCT 189
 QY 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
 DB 190 AGACAGACATACACAGCTGTGCCAGACTCTGCTTAAACCAAGAACAGTATGTTAGCA 249
 QY 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
 DB 250 AGCACTGCTATGAAGACTGAGAATTGTGTGGCCAAACAAAACTTGCCAATGGCACTTCC 309
 QY 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
 DB 310 AGTATGATTGTGCCAAGCAACGAAACTCTCAGCAAGCTATGAAAAGGAAAGAACTG 369
 QY 101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
 DB 370 TGTGTCAAAATACCTTTGAGCAGTGGTCAGAGTCAGATCAAGTGGAAITTTGTGGAACATCTT 429
 QY 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrIleuLysProMetLeu 140
 DB 430 ATATCCCAATGTGTCATTACCAACATGGGCACATAAACTCGTATCTTAAACCTATGTTG 489
 QY 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
 DB 490 CAGAGAGATTTCATTAATCTCTGCCAGCTCGGGGATTCGATCATATCGCTGAGAACATT 549
 QY 161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 180
 DB 550 CTGTCTATACCTGGATGCCAAATCACTATGTGCTGTGTAACCTGTGTGCAAGGAATGGTAC 609
 QY 181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
 DB 610 CGAGTGCCTCTGATGGCATGTGTGGAGAGAGCTTATCGAGAGAAATGGTCAGGACAGAT 669
 QY 201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220
 DB 670 TCTCTGGAGAGGCGCTGGCAGAACCGAGAGGATGGGACAGTATTTATCAAAAACAAA 729
 QY 221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
 DB 730 CCTCCTGACGGGAATGCTCTCCCAACTCTTTTATAGAGCACTTTATCCTAAATATATA 789
 QY 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
 DB 790 CAAGACATTCAGACATAGATATCTAATGGAGATGTGGAGACATAGTTTACAGAGAAAT 849
 QY 261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
 DB 850 CACTGCCGAAGTGAACCAAGCAAAAGGAGTTTACTGTTTACAGTATGATGATCAGAAAAATA 909
 QY 281 ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300
 DB 910 GTAAGCGGCTCTCGAGACAACACAATCAAGATCTGGGATAAAAACACATTGGAATGCAAG 369
 QY 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
 DB 970 CGAATTCACAGGCCCATACAGGTTCAAGTCCCTCTGCTCCAGTATGATGAGAGAGTGATC 1029
 QY 321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
 DB 1030 ATAACAGAGATCATCGGATTCACGGTCAGAGTGTGGGATGTAATATACAGGTGAAATGCTA 1089
 QY 341 AsnThrLeuIleHisIscysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360

1090 AACAGTTGATTCACCATTTGTGAAGCAGTTCTGCACCTTGGTTTCAATATGCGCATG 1149
 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
 1150 GTGACCTGCTCCAAAGATCGTTCATTCTGTATGGGATATGGCCTCCCAACTGACATT 1209
 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAsp 400
 1210 ACCCTCGGAGGTGCTGTGGACACCGAGCTGCTCAATGTTGTAGACTTGTATGAC 1269
 401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
 1270 AAGTACATTGTTCTGCATCTGGGATAGAACTATAAAGTATGGAAACAACAGTACTTGT 1329
 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
 1330 GAATTTGAGGACCTTAATATGGACAAACAGAGCAATTCCTGTTTGAGTACAGGGAC 1389
 441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGlyCysGly 460
 1390 AGGCTGTAGTGTAGTGTCTCATCTGCACACACTATCAGATTATGGGACATAGATGTGT 1449
 461 AlaCysLeuArgValLeuGluGlyHisGluLeuValArgCysIleArgPheAspAsn 480
 1450 GCATGTTTACAGATGTTAGAGGCCATGAGGAATTGGTGGTGTATTTCGATTGTATAAC 1509
 481 LysArgIleValSerGlyValAlaTrpAspGlyLysIleLysValTrpAspLeuValAla 500
 1510 AAGAGGATAGTCAGTGGGCTATGATGGAAATTAAGTGTGGGATCTTGTGGTGTCT 1569
 501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
 1570 TTGACCCCCGTGCTCTCGAGGACACTCTGTCTACGGACCCCTTGTGGAGCATTCGGA 1629
 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
 1630 AGAGTTTTCGACTACAGTTTGATGAATTCAGATTGTCAGTAGTTCAATGATGACACA 1689
 541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
 1690 ATCTCATCTGGGACTTCTTAATGATCCAGCTGCCCAAGCTGAACCCCGGCTCCCT 1749
 561 SerArgThrTyrThrIleSerArg 569
 1750 TCTCGAATACACCTACATCTCCAGA 1776

6
0

C84610 standard; DNA; 2151 BP.

C84610;

-APR-2001 (first entry)

man ZF11 protein encoding DNA.

phase kinase associated protein; SKP1; SKP2; SKP2-like protein; ZF;
 L-1; cullin; CDC53; p27; cyclin E; Max; Mad; c-Myc; MDM2; p53; Bax;
 3; Bcl-2; tumour; cytosolic; ds.

no sapiens.

200075184-A1.

-DEC-2000.

-JUN-2000; 2000WO-US15449.

-JUN-1999; 99US-0137494.

fYA) UNIV YALE.

ang H, Tsvetkov LM, Kondo T;

XX WPI; 2001-061703/07.
 DR P-ESDB; AAB48298.
 XX
 PT Modulating polypeptide levels in a cell, diagnosing and treating tumor,
 involves altering levels of proteins such as S-phase kinase associated
 PT proteins 1, 2 and cullin/CDC53 proteins -
 XX
 PS Examples; Page 129-130; 162pp; English.
 XX
 CC The invention relates to methods of altering the polypeptide levels in a
 CC cell, using proteins selected from S-phase kinase associated proteins 1
 CC and 2 (SKP1, SKP2), SKP2-like proteins (ZF) and CUL-1 (a member of the
 CC cullin/ CDC53 family of proteins). The method is useful for altering the
 CC level of p27, cyclin E, Max, Mad, c-Myc, MDM2, p53, Bax, Bad or Bcl-2
 CC polypeptide in a cell. SKP2 and SKP2-like protein levels are useful for
 CC detecting tumours, and in monitoring tumor treatment in a mammal. Agents
 CC that modulate interactions between SKP and target proteins are useful for
 CC treating tumours.
 XX
 SQ Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;

Alignment Scores:

Pred. No.: 5,66e-303 Length: 2151
 Score: 3034.00 Matches: 569
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-09-601-168B-2 (1-569) x AAC84610 (1-2151)

Qy 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
 Db 70 ATGACCCCGCCGCGCGGTGCTGCAAGAGAGGCACTCAGTTTATGATTCCTCAGAG 129
 Qy 21 ArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLysAsnSerLeu 40
 Db 130 AGAAGAACACTGTTAATATGCGAAACCCCTAGGAAGTAAATACCAGAGAGAAATTCATT 189
 Qy 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
 Db 190 AGACAGACATACACACAGCTGTGCCAGACTCTGCTTAAACCAAGAACACATATGTTAGCA 249
 Qy 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
 Db 250 AGCACTGCTATGAAGACTGAGAAATTGTGTGGCCAAACAAACTTGCCTAATGCGACTTCC 309
 Qy 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
 Db 310 AGTATGATTGTGCCAAGCAACGGAACCTCTAGCAAGCTATGAAAGGAGAAAGGAACTG 369
 Qy 101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
 Db 370 TGTGTCAAATACATTGAGCAGTGTGTGAGATCAAGTCAAGTGAATTTGTGGAACATCTT 429
 Qy 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
 Db 430 ATATCCCAATGTGTCATTACCAACATGGGCACTAAACTCGTATCTTTAAACCTAAGTTG 489
 Qy 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
 Db 490 CAGAGAGATTTTCATNACTGCTCTGCCAGCTCGGGGATTGATCATATCGCTGAGAACAATT 549
 Qy 161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 180
 Db 550 CTGTCTACCTGATGCAAAATCATTATGTGCTGCTGAACCTGTGTGCAAGGAATGGTAC 609
 Qy 181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
 Db 610 CGAGTGACCTCTGATGGCATGCTGTGGAGAGAACTTATCGAGAGAAATGGTCAGACAGAT 669
 Qy 201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220

```
670 TCTCTGTGGAGAGGCTGGCAGAACGAAGAGGATGGGACAGTATTTATTCAAAACAAA 729
|||||
221 PrpProAspGlyAsnAlaProProAsnSerPheTyArgAlaLeuTyProLysIleIle 240
|||||
730 CTTCTGTACGGGAATGCTCTCCCACTCTTTTATAGAGCACTTATCTTAATTTATA 789
|||||
241 GluAspIleGluThrIleGluSerAsnThrArgCysGlyArgHisSerLeuGlnArgIle 260
|||||
790 CAAGACATTGAGACAATAGAAATCTAATTCGAGATGTGGAAGACATAGTTTACAGAGAATT 849
|||||
261 HisCysArgSerGluThrSerLysGlyValTyCysLeuGlnTyRaspGlnLysIle 280
|||||
850 CACTGCCGAAGTGAACAACGAAGAGATTTACTGTTTACAGTATGATGATCAGAAAAATA 909
|||||
281 ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300
|||||
910 GTAAGGGGCTTCGAGACACACAAATCAAGATCTGGGATAAAAAACACATTGGAATGCAAG 969
|||||
301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyRaspGluArgValIle 320
|||||
970 CGAATTCTCAGCGCCATACAGGTTGAGTCTCTCTGTCCTCAGTATGATGAGAGATGATC 1029
|||||
321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
|||||
1030 ATACAGGATCATCGGATTCACGGTCAGAGTGTGGATGTAAATACAGGTGAATGCTA 1089
|||||
341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
|||||
1090 AACACGTTGATTACCACTTTGTGAAGCAGTTCTGCACTTGGTTTCAATTAATGGCATGATG 1149
|||||
361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
|||||
1150 GTACCTGCTCCAAAGATGTTCCATTGCTGTATGGATATGGCTTCCCACTGACATT 1209
|||||
381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
|||||
1210 ACCCTCCGAGGGTGTGTCGACACCGAGTGTCTCAATGTTGTAGACTTTGATGAC 1269
|||||
401 LysTyrlleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
|||||
1270 AAGTACATTGTTCTGCATCTGGGGATAGAACTATTAAGGTATGGAACCAAGTACTTGT 1329
|||||
421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyArgAsp 440
|||||
1330 GAATTTGTAGGACCTTAATGACACAACACGAGCATTCCTGTTGTGACGTACAGGAC 1389
|||||
441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
|||||
1390 AGGCTGGTAGTGTAGTGGCTCATCTGACACACACTATCAGATTATGGGACATAGAAATGTGT 1449
|||||
461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
|||||
1450 GCATGTTTACAGTGTGTAGAGGCCATAGAGAAATGGTGGTGTGTTATTCGATTGTAAC 1509
|||||
481 LysArgIleValSerGlyAlaTyAspGlyLysIleLysValTrpAspLeuValAlaAla 500
|||||
1510 AAGAGGATAGTCACTGGGGCTATGATGGAATAATTAAGTGTGGATCTTGTGGCTGCT 1569
|||||
501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
|||||
1570 TTGACACCCCGTGTCTCTGAGGAGACACTCTGTCTACGGACCCCTTGTGAGCATTTCCGGA 1629
|||||
521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
|||||
1630 AGAGTTTTCAGCTACAGTTTGTAGTTCAGATTTCAGTASITTCATGATGACACA 1689
|||||
541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProArgSerPro 560
|||||
1690 ATCTCATCTGGGACTTCTTAATGATCCAGCTGCCCAAGCTGAACCCCGGCTTCCCT 1749
|||||
561 SerArgThrTyThrTrpIleSerArg 569
|||||
```

Db 1750 TCTCGAACATACACCTACATCTCCAGA 1776

RESULT 7

AAL41041

ID AAL41041 standard; cDNA; 2151 BP.

XX AAL41041;

AC AAL41041;

XX 11-OCT-2002 (first entry)

XX cDNA of Human F-box protein FBPI SEQ ID No 1.

XX Cytostatic; immunomodulator; cardiant; antiinflammatory; antimicrobial;

KW proliferative; differentiative disorder; Skp2; F-box protein; cancer;

KW ubiquitin ligase; breast cancer; prostate cancer; ovarian cancer;

KW small cell lung carcinoma; immune disorder; parathyroid adenoma; FBP;

KW inflammatory disorder; lymphoma; major opportunistic infection;

KW certain cardiovascular disease; human; gene; ss.

XX Homo sapiens.

OS

XX WO200255665-A2.

FN

XX 18-JUL-2002.

PD

XX 07-JAN-2002; 2002WO-US00311.

PF

XX 05-JAN-2001; 2001US-260179P.

FR

XX (UJNY) UNIV NEW YORK STATE.

PA

XX Pagano M;

PI

XX WPI; 2002-599665/64.

DR

XX P-PSDB; AAO22446.

DR

Screening compounds for treating proliferative disorders, e.g. breast cancer or prostate cancer, infections or immune disorders, comprises detecting a change in the activity of Skp2 with either p27 or Cks1 -

XX Disclosure; Fig 3; 246pp; English.

XX The invention relates to screening compounds useful for the treatment of proliferative or differentiative disorders comprising detecting a change in the activity of Skp2 (F-box protein). The method is useful for screening compounds for the treatment of proliferative or differentiative disorders, particularly cancer. These compounds include small molecules, or compounds or derivatives or analogues of the new ubiquitin ligases. The compounds are useful for treating diseases such as cancer (e.g. breast cancer, prostate cancer or ovarian cancer, lymphoma, small cell lung carcinoma or parathyroid adenomas), major opportunistic infections, immune disorders, certain cardiovascular diseases or inflammatory disorders. This polynucleotide sequence represents the cDNA encoding an F-box protein (FBP) of the invention.

XX

SQ Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;

Alignment Scores:

Pred. No.: 5,66e-303 Length: 2151

Score: 3034.00 Matches: 569

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 24 Gaps: 0

US-09-601-168B-2 (1-569) x AAL41041 (1-2151)

QY 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20

Db 70 ATGACCCGCCGAGCGGCTGCTGCAAGAGAGGCACTCAAGTTTATGAATTCCTCAGAG 129

QY 21 ArgGluAspCysAsnAsnGlyGluProArgLysIleIleProGluLysAsnSerLeu 40

```

130 AGAGAGACTGTAATAATGGGAAACCCCTAGGAGATATAACGAGAGAATTCACCTT 189
41 ArgGlnThrTyrPheSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
190 AGACAGACATACACACAGCTGTGCAGACTCTGCTTAAACCAAGAAACACATATGTTAGCA 249
61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
250 AGCACTGCTATGAAGACTGAGAATTGTGTGCCCAAAACAACTTGCCAAATGGCACTTCC 309
81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGlnLysGluLeu 100
310 AGTATGATTGTGCCCAAGCAACGGAACCTCTCGACAGCTATGAAAGAAAGAAAGCACTG 369
101 CysValLysTyrPheGluGlnTyrSerGluSerAspGlnValGluPheValGluHisLeu 120
370 TGTGTCAAAATCTTTGAGCAGTGTGTGAGAGTCAGATCAAGTGGAAATTTGTGGAACATCTT 429
121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
430 ATATCCCAAAATGTGTCAATACCAACATGGGCACATAAATCTCGTATCTTAAACCTATGTTG 489
141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
490 CAGAGAGATTTCAATACTGCTCTGCCAGCTCGGGGATTTGGATCATATCCTGAGAACATT 549
161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTyrTyr 180
550 CTGTCACTACCTGGATGCCAAATCACTATGTGTCTGCTGAACTTGTGTGCAAGGAATGGTAC 609
181 ArgValThrSerAspGlyMetLeuTyrLysLysLeuIleGluArgMetValArgThrAsp 200
610 CGAGTGACCTCTGATGGCATGCTGTGGAAGAAGCTTATCGAGAGAATGTCACGAGACAGAT 669
201 SerLeuTyrArgGlyLeuAlaGluArgGlyTyrGlyGlnTyrLeuPheLysAsnLys 220
670 TCTCTGTGGAGAGCCCTGGCAGAACGAAGAGGATGGGACAGTATTTATTCAAAACCAA 729
221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
730 CCTCCTCAGCGGAATGCTCCTCCCACTCTTTTATAGACACTTTATCTCTAAATATATA 789
241 GlnAspIleGluThrIleGluSerAsnTyrArgCysGlyArgHisSerLeuGlnArgIle 260
790 CAAGACATTGAGACATATAGATCTAATTTGAGATGTGGAAGACATAGTTTACAGACAA 849
261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
850 CACTGCCGAAGTGAACCAAGCAAGGAGTTTACTGTTTACAGTATGATGATCAGAAATA 909
281 ValSerGlyLeuArgAspAsnThrIleLysIleTyrAspLysAsnThrLeuGluCysLys 300
910 GTRAAGGCCCTTCAGACACACACAAATCAAGATCTGGATATAAACACATTTGNAATGCAG 969
301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
970 CGAATCTTCACAGGCCCATACAGGTTTCACTGCTCTGCTCCAGTATGATGAGAGATGATC 1029
321 IleThrGlySerSerAspSerThrValArgValTyrAspValAsnThrGlyGluMetLeu 340
1030 ATACAGAGATCATCGGAATCCAGGTGAGAGTGGGATGTAATACAGGTGAATGCTA 1089
341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
1090 AACAGTTGATTCACTTCACTTGTGAAGCAGTTCTGCACCTTGGCTTCAATATGTCATGATG 1149
361 ValThrCysSerLysAspArgSerIleAlaValTyrPhePheMetAlaSerProThrAspIle 380
1150 GTGACCTGCTCCAAAGATCGTTCATTTGCTGTATGGGATATGGCTCCCACTGACATT 1209
381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
1210 ACCCTCCGAGGGTGTGCTGGACCGAGCTGCTGTCATGTTGTAGACTTTGTATGAC 1269
401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTyrAsnThrSerThrCys 420
1270 AAGTACATTGTTTCTGCATCTGGGATAGAACTATAAAGGTATGGAACACAACTACTTGT 1329
421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
1330 GAATTTGTAAGGACCTTAATAGCACAAACGAGGCATTTGCCCTGTTTCAGTACAGGGAC 1389
441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTyrAspIleGluCysGly 460
1390 AGCTCGTATGTAGTGGCTCATCTGACAAACACTATCAGATTATGGGACATAGAATGTGT 1449
461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
1450 GCATGTTTACGAGTGTAGAGGCCATGAGGAATTTGGTCCGTTGTATTTCGATTGATTAAC 1509
481 LysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTyrAspLeuValAlaAla 500
1510 AAGAGGATAGTCAGTGGGCCCTATGATGGAATAATTAAGTGTGGGATCTTTGTGGCTGCT 1569
501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
1570 TTGGACCCCGTCTCTCTGAGGACACTCTCTTACGACCCCTTGTGAGACATTTCCGGA 1629
521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
1630 AGAGTTTTCGACTACAGATTGATGAATTCAGATTGTTCAGTAGTTCAATGATGACACA 1689
541 IleLeuIleTyrAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
1690 ATCTCATCTGGGACTTCTTAAATGATCCAGCTGCCAAGCTGAACCCCGTTCCTTCCCT 1749
561 SerArgThrTyrThrTyrIleSerArg 569
1750 TCTCGAACATACACCTTACATCTCCAGA 1776

RESULT 8
ABS51009
ID ABS51009 standard; cDNA; 2151 BP.
XX
AC ABS51009;
XX
DT 21-OCT-2002 (first entry)
XX
DE Human cDNA encoding bait protein beta-TrCP1.
XX
KW Human; ss; gene; yeast two-hybrid assay; adipocyte; bait protein; NIDDM;
non-insulin diabetes mellitus; obesity; selected interacting domain;
XX
KW SID; protein-protein interaction map; PIW; anorectic; metabolic disorder.
XX
OS Homo sapiens.
XX
PN WO200253726-A2.
XX
PD 11-JUL-2002.
XX
PF 28-DEC-2001; 2001WO-BP15423.
XX
PR 02-JAN-2001; 2001US-259377P.
XX
PA (HYBR-) HYBRIGENICS.
XX
PI (CNRS) CENT NAT RECH SCI.
XX
PI Legrain P, Marullo S, Jockers R;
XX
DR WPI; 2002-583612/62.
DR P-PSDB; ABG69473.
XX
PT Novel complex of protein-protein interactions in adipocyte cells for
identifying compounds that modulate the protein-protein interactions
PT and useful for treating obesity and metabolic disorders.
XX

```


claim 2; Page -; 125pp; English.

The invention relates to a complex of protein-protein interactions forming a protein-protein interaction map, (PIM) in adipocyte cells as defined in the specification, or polynucleotides in adipocytes encoding for the polypeptides. Also included are a recombinant cell expressing the interacting polypeptides and a method of selecting a modulating compound in adipocyte cells, by culturing a recombinant host cell on a selective medium containing a modulating compound and a reporter gene the expression of which is toxic for the recombinant host cell which is transformed with two vectors, where the first vector comprises a polynucleotide encoding a first hybrid polypeptide and DNA binding domain and the second vector comprising a polynucleotide encoding a second hybrid polypeptide and an activating domain that activates the basic reporter gene, when the first and second hybrid polypeptides interact and selecting the modulating compound which inhibits the growth of the recombinant host cell (i.e. using the yeast two-hybrid system). The complexes are useful for identifying compounds that modulate protein-protein interactions and useful for treating obesity and metabolic disorders e.g. non-insulin dependent diabetes mellitus, and the compound isolated by the method is useful for treating and preventing obesity or metabolic diseases. The interactions between proteins of the complex further define a set of selected interacting proteins, SID. The present sequence encodes a member of the protein complex of the invention, used as the bait protein in the yeast two-hybrid assay.

The present sequence was not displayed in the specification but is obtained from its Genbank entry by the indexer.

quence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;

nt Scores:

O.:	5.66e-303	Length:	2151
Similarity:	100.00%	Matches:	569
cal Similarity:	100.00%	Conservative:	0
atch:	100.00%	Mismatches:	0
	24	Indels:	0
		Gaps:	0

01-168B-2 (1-569) x ABSS1009 (1-2151)

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
 70 ATGGACCGCGCGAGCGGTGCTGCAAGAGAGAGCACTCAAGTTTATGAATTCCTCAGAG 129
 21 ArgGluAspCysAsnAsnGlyGluProPrArgLysIleLeuProGluLysAsnSerLeu 40
 130 AGAGAAGACTGTATATGGCGAACCCTTAGGAAGATAATACAGAGAGAAATTCACCT 189
 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
 190 AGACAGACATACACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTATGTTAGCA 249
 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLysLeuAlaAsnGlyThrSer 80
 250 AGCACTGCTATGAAGACTGAAGATGTGTGGCCAAAACAAACTTCCCAATGSCACTTCC 309
 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
 310 AGTATGATTGTGCCAAGCAGAAACTCTCAGCAGCTATGAAAGGAAAGGAAGCTG 369
 101 CysValLysTyrPheGluGlnThrSerGluSerAspGlnValGluPheValGluHisLeu 120
 370 TGTGTCAAAATCTTTGAGCAGTGGTTCAGAGTCAGATCAAGTGGAAATTTGTGGAACTCTT 429
 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
 430 ATATCCCAATGTGTCAATACCAACATGGGCACATAAATCGTATCTTAAACCTATGTG 489
 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
 490 CAGAGAGATTTTCATAACTGCTCTGCCAGCTGGGGATTGGATCATATCGCTGAGACATT 549

QY	161	LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr	180
Db	550	CTGTACATACCTGATGCAAAATCACTATGTGCTGCTGAACTTGTCGAAGGAATGCTAC	609
QY	181	ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValargThrAsp	200
Db	610	CGAGTGACCTCTGATGCGCATGCTGTGGAAGAGCTTATCGAGAGAAATGTCAGGACAGAT	669
QY	201	SerLeuTrpArgGlyLeuAlaGluArgGlyTrpGlyGlnTyrLeuPheLysAsnLys	220
Db	670	TCTCTGTGGAGAGCGCTGGCAGAACCGAAGAGGATGGGACAGTATTTATCAAAACAAA	729
QY	221	ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle	240
Db	730	CCTCCTGACGGGAATGCTCTCCCAACTCTTTATAGAGCACTTTATCTTAAATATA	789
QY	241	GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle	260
Db	790	CAAGACATTGAGACATAGATCTAATTGGAGATGTGGAGACATAGTTTACAGAGATT	849
QY	261	HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle	280
Db	850	CACCTGCCGAAGTGAACCAAGAAAGGAGTTTACTGTTTACAGTATGATGATCAGAAAAATA	909
QY	281	ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys	300
Db	910	GTAAGCGGCTTCGAGACACACAAATCAAGATCTGGGATAAACAACATTTGGAATGCAAG	969
QY	301	ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle	320
Db	970	CGAATTCACAGGCCCATACAGGTTCAAGTCTCTGCTCCAGTATGATGAGAGAGTGATC	1029
QY	321	IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu	340
Db	1030	ATAACAGGATCATCGGATTCACCGGTGAGAGTGTGGGATGTAATACAGGTGAAATGCTA	1089
QY	341	AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet	360
Db	1090	AACAGGTGATTCCACCATTTGGAAGCAGTTCTGCACTTTCATTAATGSCATGATG	1149
QY	361	ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle	380
Db	1150	GTGACCTGCTCCAAAGATCGTTCCATTGCTGTATGGGATATGGCCTCCCAACTGCATT	1209
QY	381	ThrLeuArgArgValLeuValGlyHisArgAlaValAsnValValAspPheAspAsp	400
Db	1210	ACCTCCGGAGGTGCTGCTCGGACACCGAGCTGCTGTAATGTTGTAGACTTTGATGAC	1269
QY	401	LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys	420
Db	1270	AAGTACATTGTTCTGCACTCGGGATAGAACTATAAAGGTATGGAACACACAGTACTGT	1329
QY	421	GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp	440
Db	1330	GAAATTTGATAGGACCTTAAATGGACACAAACGAGGCAATTCCTGTTTTCAGTACAGGAC	1389
QY	441	ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly	460
Db	1390	AGGCTGTGTAGTGTGCTCATCTGCAACACACTATCAGATTATGGGACATAGATCTGGT	1449
QY	461	AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn	480
Db	1450	GCATGTTTACAGTGTGTAGAGGCCCATGAGGAATTCGTGCTGTATTCGATTGTGATAAC	1509
QY	481	LysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAla	500
Db	1510	AAGAGGATAGTCACTGGGCGCTATGATGGNAAATTAAGTGTGGGATCTTGTGGCTGCT	1569
QY	501	LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly	520
Db	1570	TTGACCCCGTGTCTCTGCGAGGACACTCTGCTACGAGACCCCTTGTGGAGCATTCGGA	1629
QY	521	ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerSerHisAspAspThr	540

|||||
1630 AGAGTTTTCGACTACAGTTTGAATTCAGATTGTCAGTAGTTTCACATGATGACACA 1689
|||||
541 lileullelfrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProArgSerPro 560
|||||
1690 ATCTCTCATCTGGGACTTCTCTAAATGATCCAGCTGCCCAAGCTGAACCCCGCTTCCCT 1749
|||||

561 SerArgThrTyrThrTyrIleSerArg 569
|||||
1750 TCTGCACATACACTACATCTCCAGA 1776
|||||

9
33
1229233 standard; cDNA; 2419 BP.

1229233;

-FEB-2000 (first entry)

man cell signalling protein-12 encoding cDNA.

ll signalling protein-12; CSIGP-12; cell proliferation;
flammatory disorder; cirrhosis; cancer; hepatitis; ss.
teriosclerosis; Addison's disease; multiple sclerosis; ss.

mo sapiens.

Y Location/Qualifiers
S 70..1779
/*tag= a
/product= "Cell Signalling Protein-12"

9958558-A2.

-NOV-1999.

99NOV-1999; 99NO-US10567.

99NOV-1998; 98US-0085343.
-AUG-1998; 98US-0098010.

NCY-) INCYTE PHARM INC.

ndman O, Hillman JL, Lal P, Yue H, Tang YT, Patterson C;
Jghn MR, Yang J;

I; 2000-086432/07.
PSDB; AAY44249.

nan cell signaling proteins useful for, e.g. diagnosing cell
proliferative and inflammatory disorders

aim 9; Page 87-88; 90pp; English.

a present sequence is a cDNA obtained from Incyte clone 3239149 of
-AUCT01 library. It encodes cell signalling protein-12 (CSIGP-12). It
expressed in musculo-skeletal, gastrointestinal and nervous
issues. Fragments of CSIGP encoding nucleic acid can be used as
proliferation probe for detecting CSIGP related sequences or allelic
variants. Recombinant CSIGP can be produced in host cells by transforming
am with genetically engineered vectors. Agonists or antagonists can be
used in the treatment of cell proliferative and inflammatory disorders
associated with decreased or increased CSIGP expression. CSIGP is used in
diagnosis, prevention and treatment of cell proliferative disorders
arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory
disorders like AIDS, Addison's disease, multiple sclerosis, etc.

quence 2419 BP; 671 A; 531 C; 625 G; 590 T; 2 other;

it Scores:

6.75e-303 Length: 2419
3034.00 Matches: 569
Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0
US-09-601-168B-2 (1-569) x AA229233 (1-2419)
Qy 1 MerAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
Db 70 ATGGACCCGCCGAGGGCTGCTCAAGAGAGAGGCACTCAAGTTTATGAATTCCTCAGAG 129
Qy 21 ArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLysAsnSerLeu 40
Db 130 AGAGAGACTGTAAATAATGCGAACCCTCTAGGAAGATAATACAGAGAGAAATTCACCT 189
Qy 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
Db 190 AGACAGACATACACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTATGTTAGCA 249
Qy 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
Db 250 AGCACTGCTATGAGACTGAGAAATGCTGTGGCCAAACAAACTTGCCAATGGCACTTCC 309
Qy 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
Db 310 AGTATGATTGTGCCCAAGCAACGAAACTCTCAGCAAGCTATGAAAGGAAAGGAAGCTG 369
Qy 101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
Db 370 TGTGTCMAATACCTTTCAGCAGTGTGTCAGAGTCAGATCAAGTGGAAATTTGTGGAACATCT 429
Qy 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
Db 430 ATATCCCAATGTGTCTATTACCAACATGGGCACATAAACTCGTATCTTAAACCTATGTTG 489
Qy 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGlnAsnIle 160
Db 490 CAGAGAGATTTCATAACTGCTCTGCCAGCTCGGGATGGATCATATTGCTCAGAACATT 549
Qy 161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 180
Db 550 CTGTACATACCTGGATGCCAAATCCTATGCTGTCTGTAACCTGTGTGCAAGGAATGGTAC 609
Qy 181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
Db 610 CGAGTGACCTCTCATGSCATGCTGTGGAAGAAAGCTTATCGAGAGAAATGGTCAGACAGAT 669
Qy 201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220
Db 670 TCTCTGTGGAGAGGCTTGGCAGNACGAGAGAGATGGGACAGATATTTATCAAAAACAAA 729
Qy 221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
Db 730 CCTCTGTACGGGAATGCTCTCTCCAACTCTTTTATAGAGCACTTTATCTTAAAAATTATA 789
Qy 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
Db 790 CAAGACATTGAGACATAGAACTAATTCGAGATGTCGAGACATAGTTTACAGAGAAATT 849
Qy 261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
Db 850 CACTGCCGAAGTGAACCAAGCAAGAGAGTTTACTGTTTACAGTATGATGATCAGAAAATA 909
Qy 281 ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300
Db 910 GTTAGGGGCTTTCGAGACCAACCAATCAAGATCTGGGATAAAAAACACATTGGAATGCAAG 969
Qy 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
Db 970 CGAATTCACAGGCCATACAGGTTTCAGTCTCTCTCTCCAGTATGATGAGAGAGTGATC 1029
Qy 321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
Db 1030 ATAACAGGATCATCGGATTCACGGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGT 1089

341 AsnThrLeuIleHisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
 1090 AACACGTTGATTACCATTTGTGAAGCAGTTCGCACTTCGGTTTCAATTAATGGCATGATG 1149
 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
 1150 GTGACCTGCTCCAAAGATGCTTCATGCTGTGTATGGATATGGCTCCCACTGACAT 1209
 381 ThrLeuArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
 1210 ACCCTCGGAGGTGCTGTCGACACCGAGCTGCTCAATGTTGTAGACTTTGATGAC 1269
 401 LysTyrlleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
 1270 AAGTACATTTGTTCTGCATCTGGGATAGAACTATTAAGGTATGGAACAACAAGTACTTGT 1329
 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTrpArgAsp 440
 1330 GAATTTGTAGGACCTTAATGACACACACGAGCATTCCTGTTTGCAGTACAGGGAC 1389
 441 ArgLeuValValSerGlySerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
 1390 AGGCTGTGTAGTGTGCTCATCTGACAACTATCAGATTATGGGACATAGAAATGCTGT 1449
 461 AlaCysLeuArgValLeuGluGlyHisGluLeuValArgCysIleArgPheAspAsn 480
 1450 GCATGTTTACAGTGTAGAGGCCATGAGGAATGTCGTTGTATTCATTTGATTAAC 1509
 481 LysArgIleValSerGlyAlaTrpAspGlyLysIleLysValTrpAspLeuValAla 500
 1510 AAGAGGATAGTCAGTGGGGCTATGATGAAAAATTAAGATGTGGGATCTTGTGGCTGT 1569
 501 LeuAspProGlnAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
 1570 TTGACCCCGTCTCTCGAGGACACTGTCTACGGACCCCTGTGAGCATTCGGA 1629
 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
 1630 AGAGTTTTCGACTACAGTTTGTATGAATTCAGATTGTCAAGTATTCATGATGACACA 1689
 541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
 1690 ATCTCATCTGGGACTTCCTAAATGATTCAGCTGCCAAGCTGAACCCCGTTCCCT 1749
 561 SerArgThrTrpThrTrpIleSerArg 569
 1750 TCTCGACATACACCTACATCTCCAGA 1776

10

9

490079 standard; cDNA; 3220 BP.

490079;

-OCT-2001 (first entry)

nan bone marrow cDNA, SEQ ID NO: 323.

nan; bone marrow; antiinflammatory; cytostatic; neuroprotective;
 viral; antibacterial; antifungal; anti-HIV; haemostatic;
 immunosuppressive; gene therapy; cytokine cell proliferation;
 differentiation modulator; immune disorder; infection; cancer;
 nan immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss.

no sapiens.

200153453-A2.

JUL-2001.

DEC-2000; 2000MO-US34960.

PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 30-NOV-2000; 2000US-0250583.
 XX (HYSE-) HYSEQ INC.
 XX
 PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
 PI Ren F, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Drmanac RT;
 DR WPI; 2001-488707/53.
 DR P-PSDB; AAM00960.

Novel bone-marrow-expressed polynucleotides and polypeptides, useful
 for treating e.g. cancer and immune deficiency disorders -

Claim 1; Page 428; 648pp; English.

The present sequence is one of 251 novel human polynucleotides
 expressed in the bone marrow. The polynucleotide and the
 polypeptide encoded by it are useful in the treatment of various
 immune deficiencies and disorders. The deficiencies and disorders may
 be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
 infection, or may result from an autoimmune disorder, a coagulation
 disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
 suppression of an inflammatory response or treatment of a nervous
 system disorder such as Alzheimer's disease. Detection of the presence
 or increased expression of the polynucleotide or the protein it
 encodes is useful for the diagnosis and/or prognosis of one
 or more types of cancer. The polynucleotide and polypeptide can be
 used as nutritional sources or supplements and in the screening of
 chemical compounds as potential drugs.

Sequence 3220 BP; 878 A; 725 C; 770 G; 847 T; 0 other;

Alignment Scores:

Pred. No.:	5,46e-302	Length:	3220
Score:	3027.00	Matches:	567
Percent Similarity:	99.82%	Conservative:	1
Best Local Similarity:	99.65%	Mismatches:	1
Query Match:	99.77%	Indels:	0
DB:	22	Gaps:	0

US-09-601-168B-2 (1-569) x AAH90079 (1-3220)

Qy 1 MetAspProAlaGluAlaValLeuGlnGlyLysAlaLeuLysPheMetAsnSerSerGlu 20
 Db 128 ATGGACCGGTGCGGGGGTGTGCAAGAGAGAGGCACTCAAGTTTATGAATTCCTCAGAG 187
 Qy 21 ArgGluAspCysAsnAsnGlyGluProArgLysIleIleProGluLysAsnSerLeu 40
 Db 188 AGAGAAGACTGTATTAATGCGGAACCCCTAGAGAGATAATACCAGAGAAGAAATTCGCTT 247
 Qy 41 ArgGlnThrTrpAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
 Db 248 AGACAGACATACCAACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTATGTTTAGCA 307
 Qy 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
 Db 308 AGCACTGCTATGAAGACTGAGAATTCGTGGCCAAACAAAACTTGCCAATGGCACTTCC 367
 Qy 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrgLysGluLysGluLeu 100
 Db 368 AGTATGATTGTGCCCAAGCAACGAAACTCTCAGCAAGCTATCAAAAGGAAAGGAACTG 427
 Qy 101 CysValLysTyrrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
 Db 428 TGTCGTAATACATTTGAGCAGTGTGTCAGAGTCAGATCAAGTGGGANTTTGTGGAACATCTT 487

```

121 IleSerGlnMetCysHisTyrClnHisIleSclHisIleAsnSerTyrLeuLysProMetLeu 140
|||||
488 ATATCCCAAAATGTGTCAATTACCAATCGGCACATAAATCGTATCTTAAACCTATGTTG 547
|||||
141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
|||||
548 CAGAGAGATTCATCAACTGCTCTGCCAGCTCGGGATGGATCATATTCGTGAGACATT 607
|||||
161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTyrTyr 180
|||||
608 CTGTCTATACCTGGATGCCAAATCAATATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 667
|||||
181 ArgValThrSerAspGlyMetLeuTyrLysIleLeuIleGluArgMetValArgThrAsp 200
|||||
668 CGAGTGACCTCTGATGGCATGCTGTGGAGAGCTTATCGAGAGATGGTCAGACAGAT 727
|||||
201 SerLeuTyrArgGlyLeuAlaGluArgGlyTyrGlyGlnTyrLeuPheLysAsnLys 220
|||||
728 TCTCTGTGGAGAGGCTGCGAGACGAAGAGGATGGGACAGATATTTATTCAAAACAAA 787
|||||
221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
|||||
788 CCTCTGACGGGAATGCTCTCCCAACTCTTTTATAGAGCACTTTATCTAAATATA 847
|||||
241 GlnAspIleGluThrIleGluSerAsnTyrArgCysGlyArgHisSerLeuGlnArgIle 260
|||||
848 CAAGACATTTGAGACATAGATCTAATTTGGAGATGGAGACATAGCTTTACAGAGATT 907
|||||
261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
|||||
908 CACTCCGGAAGTGAACGAAGCAAGAGGAGTCTACTGTTTACAGTATGATGATCGAATA 967
|||||
281 ValSerGlyLeuArgAspAsnThrIleLysIleTyrAspLysAsnThrLeuGluCysLys 300
|||||
968 GTACGCGGCTTCGAGACACACATCAAGATCTGGGNTAAACACATTTGGATGCAAG 1027
|||||
301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
|||||
1028 CGAATCTCACAGGCATACAGGTTTCAGTCTCTGCTCCAGTATGATGAGAGATGATC 1087
|||||
321 IleThrGlySerSerAspSerThrValArgValTyrAspValAsnThrGlyGluMetLeu 340
|||||
1088 ATACAGGATCATCGAATTCACGCTCAGAGTGGGATGTAATACAGGTGAATGCTA 1147
|||||
341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
|||||
1148 AACACGTTGATTCACCATTTGAAGCAGTTCTGCACCTGCGTTTCAATATGCGCATG 1207
|||||
361 ValThrCysSerLysAspArgSerIleAlaValTyrAspMetAlaSerProThrAspIle 380
|||||
1208 GTGACCTGCTCCAAAGATCGTTCCATTGCTGTATGGGATATGGCTCCCAACTGACATT 1267
|||||
381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValAspPheAspAsp 400
|||||
1268 ACCCTCCGAGGGTCTGTCGACACCGAGCTGCTGTCATGTTGTAGACTTTGATGAC 1327
|||||
401 LysTyrIleValSerAlaSerClyAspArgThrIleLysValTyrAsnThrSerThrCys 420
|||||
1328 AAGTACATGTTTCTGCAATCTGGGATAGAACTATAAGGTTATGGAAACACAAGTACTGT 1387
|||||
421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
|||||
1388 GAATTTGTAAGGACCTTAATAGACACAAACGAGCATTTGCTGTTTGCAGTACAGGGAC 1447
|||||
441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTyrAspIleGluCysGly 460
|||||
1448 AGGCTGGTAGTGGTGGTCTCTCTCACACACTATCAGATTATGGGACATAGATGGT 1507
|||||
461 AlaCysLeuArgValLeuGluClyHisGluLeuValArgCysIleArgPheAspAsn 480
|||||
1508 GATGTTTACGAGTGTTAAGAGGCCATGAGGAATGGTGGCTGTTGATTCGATTGATAAC 1567
|||||

```

RESULT 11

```

AAK51715
ID AAK51715 standard; cDNA; 2285 BP.
XX
AC AAK51715;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 260.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US040398.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0661561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
(PHYSE-) HYSEQ INC.
XX
PA
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI; 2001-478283/51.
DR P-PSDB; AAM78582.
XX
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
PS Claim 1; Page 1177-1180; 6221pp; English.
XX
XX
CC The invention relates to polynucleotides [AAK51456-AAK53435] and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating

```

tivity, tissue growth factor activity, immunomodulatory activity and
tissue inhibitor activity and may be useful in the diagnosis and/or
treatment of cancer, leukaemia, nervous system disorders, arthritis and
inflammation.
Reference: for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
(AAK80020) are omitted as the relevant pages from the sequence listing
are missing at the time of publication.

sequence 2285 BP; 646 A; 492 C; 573 G; 574 T; 0 other;

```

nl Scores:      4.94e-300      length:      2385
co. :           300e.00      Matches:      569
              94.05%      Similarity:      0
              94.05%      cal Similarity:      0
              99.08%      atCh:      36
              22          Gaps:      1

```

01-168B-2 (1-569) x AAK51715 (1-2285)

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLys----- 14

127 ATGGACCCGCGAGGCGTGCTGCAAGAGAGGCACTCAAGTTATGTGCTCTATGCC 186

14-----14

187 AGGTCTCTGTGGCTGGGCTGCTCCAGCCTGGCGGACAGCATGCCCTTCGCTGGGATGCCCTG 246

15 -----PheMetAsnSerSerGluArgGluAspCys 24

247 TATAACCCAGGACTGGCGCACTCACAGCTTTCATGAATTCTCAGAGAGAGACTGT 306

25 AsnAsnGlyGluProProArgLysIleIleProGluLysAsnSerLeuArgGlnThrTyr 44

307 AATAATGGCGAACCCCTAGGAAGATAATACCAGAGAAGAAATTCGCTTAGACAGACATAC 366

45 AsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAlaSerThrAlaMet 64

367 AACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTATGTTTAGCAAGCACTGCTATG 426

65 LysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSerSerMetIleVal 84

427 AAGACTGAGAAATTGTGTGGCCAAACAAACTTGCCAATGGCACTTCCAGTATGATTGTG 486

85 ProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeuCysValLysTyr 104

487 CCCAAGCAACGGAACTCTCAGCAAGCTATGAAAGGAAAGGAAGGAACTGTGTGTCAAATAC 546

105 pheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeuIleSerGlnMet 124

54 / TTTGAGCAGTGGTCAGAGTCAGATCAAGTGGAAATTGTGGAAACATCTTATATCCCAATG 606

123 C Y S H I S I Y I G I N H I S G I Y H I S I I E A S H S E R I Y L E U L Y S F R O M E T L E U G I N A R G A S P P H E 144

60' TGT CAT TAC CCA CACA TGG GCA CATA TTA C TCG TAT C TTA TTA C CTA TAT GTT G CAG AGA GAT T TC 668

[illegible]

GG, AATACCTGCTCAGCCTCGGGAATTGGAAICATAATGCTCGAGAACATTCGTCAIACCIG /Z8

[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

907	Db	AAATGCTCTCCCAACTCTTTTATAGAGCACTTTATCCTAAAAATTATCAAGACATTGAG	966
245	Qy	ThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSer	264
967	Db	ACAAATAGACTCTAAATGGAGATGTGGAGACATAGTTTACAGAGAATTCTACTGCCGAAGT	1026
265	Qy	GluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLeu	284
1027	Db	GAARCAAGCAAAAGAGTTTACTGCTTTACAGTATGATCATCAGAAAAATAGTAAAGCGGCGTT	1086
285	Qy	ArgAspAsnThrIleLysIleTyrAspLysAsnThrLeuGluCysLysArgIleLeuThr	304
1087	Db	CGAGACAAACAAATCAAGATCTGGGATAAAAACACATTTGGATGCAAGCGAATTCACACA	1146
305	Qy	GlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIleIleThrGlySer	324
1147	Db	GGCCATACAGGTTCAAGTCTCTGCTCCAGTATGATGAGAGAGTGAATCAAAACAGGATCA	1206
325	Qy	SerAspSerThrValArgValTyrAspValAsnThrGlyGluMetLeuAsnThrLeuIle	344
1207	Db	TCGGATTCCACGGTCAGAGTGTGGATGTAAATACAGGTCAAAATGCTAAACAGTTGAATT	1266
345	Qy	HisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSer	364
1267	Db	CACAAITGTGAAGCAGTTCTGCACCTTCGCTTCCAAATAATGGCATGATGTTGACCTGCTCC	1326
365	Qy	LysAspArgSerIleAlaValTyrAspMetAlaSerProThrAspIleThrLeuArgArg	384
1327	Db	AAAGATCGTTCCATTGCTGTATGGGATATGGCCCTCCCACTGACATTACCCCTCCGGAGG	1386
385	Qy	ValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAspLysTyrIleVal	404
1387	Db	GTGCTGTGTCGGACACCGAGCTGTGTCAAATGTTGTAGACTTTGATGACAAGTACATTGTT	1446
405	Qy	SerAlaSerGlyAspArgThrIleLysValTyrAsnThrSerThrCysGluPheValArg	424
1447	Db	TCGTGCATCTGGGATAGACTATTAAGGTATGGACACAAAGTACTTGTGTGAATTTGAAGG	1506
425	Qy	ThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAspArgLeuValVal	444
1507	Db	ACCTTAAATGGACACAAACGAGGCATTGCTGCTTTTCAGTACAGGGACAGCTGTTAGTG	1566
445	Qy	SerGlySerSerAspAsnThrIleArgLeuThrPaspIleGluCysGlyAlaCysLeuArg	464
1567	Db	AGTGGCTCATCTGACAAACACTATCAGATTATGGACATAGAATGTGGTCATGTTTACGA	1626
465	Qy	ValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsnLysArgIleVal	484
1627	Db	GTGTTAAGAGGCCATCAGGAATTTGGTGGCTGTATTTCGATTTCATAAAGAGGATAGTC	1686
485	Qy	SerGlyAlaTyrAspGlyLysIleLysValTyrAspLeuValAlaLeuAspProArg	504
1687	Db	AGTGGGGCCCTATGATGGAAAAATTAAGGTGTGGGATCTTGTGGCTGCTTTGGACCCCCGT	1746
505	Qy	AlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGlyArgValPheArg	524
1747	Db	GCTCCTCGAGGCACACTCTGCTACGACCCTTGTGGAGCATTTCGGAAGAGATTTTTCGA	1806
525	Qy	LeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThrIleLeuIleTyr	544
1807	Db	CTACAGTTTCATGAATTCAGATTGTTCAGTAGTTCATGATGACACAATTCCTCATCTGG	1866
545	Qy	AspPheLeuAsnAspProAlaAlaGlnAlaGluProArgSerProSerArgThrTyr	564
1867	Db	GACTTCTCTAAATGATCCAGCTGCCCAGAGCTGAACCCCCCGTTCCTCCGAAACATAC	1926
565	Qy	ThrTyrIleSerArg	569
1927	Db	ACCTACATCTCCAGA	1941

RESULT 12
AAA73131

A73131 standard; cDNA; 1707 BP.

A73131;

-NOV-2000 (first entry)

use ubiquitin ligase FWD1 protein encoding cDNA.

ubiquitin ligase SCF complex; F-box protein; ubiquitination; IkappaB; ta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FWD1; ne therapy; colon cancer; beta-transducin repeat containing protein; ta-TrCP; ss.

s musculus.

2000166542-A.

-JUN-2000.

-DEC-1998; 98JP-0343437.

-DEC-1998; 98JP-0343437.

AGA-) KAGAKU GIUTSU SHINKO JIGYODAN.

I; 2000-485550/43.

PSDB; AAB12812.

box protein of ubiquitin ligase SCF complex which promotes the ubiquitination of IkappaB or beta-catenin

sclosure; Fig 16; 19pp; Japanese.

e present invention describes an F-box motif protein of ubiquitin ligase SCF complex which promotes the ubiquitination of IkappaB or ta-catenin and is constituted by Skp1 protein, Cull1 protein and a mplex (SCF complex) of F-box protein containing F-box motif and WD40 repeat motif and has the amino acid sequence of 45 residues (AAB12811) one of two 569 residue sequences (AAB12812, which is mouse ubiquitin ligase FWD1 protein) and (AAB12813, which is human beta-transducin repeat containing protein (beta-TrCP)). The F-box protein can be used for e gene therapy of colon cancer by being recombined to a virus vector. e present sequence encodes the mouse ubiquitin ligase FWD1 protein, om the present invention.

quence 1707 BP; 467 A; 399 C; 453 G; 388 T; 0 other;

nt Scores:

Length:	1707
Matches:	561
Similarity:	99.12%
Conservative:	3
cal Similarity:	98.59%
at ch:	0
Indels:	0
Gaps:	0

01-168B-2 (1-569) x AAA73131 (1-1707)

1 MetAspProAlaGluValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
 1 ATGGACCCGGCAGAGCGGTGCTGCGAGGAGAAAGCGCTTAAGTTATGAATTCCTCAGAG 60
 21 ArgGluAspCysAsnAsnGlyGluProCysArgLysIlePheProGluLysAsnSerLeu 40
 61 AGAGAGACTGTAATATGCGGAACCCCTAGGAAGATTAATCCAGAGAGAAATTCACCT 120
 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
 121 AGACAGACTTACAAACAGCTGTGCCAGGCTTTGTCATAAACAAGAGACAGTATGCTAACA 180
 61 SerThAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
 181 AGCACTGCTATGAAGACTGAAATTTGTGGCCAAAGCCAAACTTGCCAAATGGCACTTCC 240

Qy 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
 Db 241 AGCATGATTTGTGCCCAAGCAGCGGAAACTCTCAGCAGCTATGAGAAAGGAGAGCTG 300
 Qy 101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
 Db 301 TGTGTCAAGTATTTTGGCAGTGGTCTGAGAGTCTGATCAAGTGGAAATTTCTAGAACACCTT 360
 Qy 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
 Db 361 ATATCCCAATGTTCTACCTACAGCATGGGCACATCACTCTCTACCTAAACCTATATGCTG 420
 Qy 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
 Db 421 CAGAGGATTTTCATAACTGCACCTGCCAGCAGCGGGTCTGACCACATCGCTGAGAACATT 480
 Qy 161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 180
 Db 481 CTGTCTACTTGGACGCCCAAGTCACTGTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 540
 Qy 181 ArgValThrSerAspGlyMetLeuTyrLysLysLeuIleGluArgMetValArgThrAsp 200
 Db 541 CGGTGACGTCCGACCGGCATGCTGTGAAAAAAGCTCATCGAGAGGATGGTTCAGGACGGAC 600
 Qy 201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTyrGlyGlnTyrLeuPheLysAsnLys 220
 Db 601 TCTCTGTGGCGAGCGCTGCCAGAGCGCAGAGGCTGGGACAGTACTTATTTCAAAAAACAAA 660
 Qy 221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
 Db 661 CTCCTCTGATGAGAAACGCTCTCTCCAACTCTTTATAGAGCGCTTTATCTAAATCATA 720
 Qy 241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
 Db 721 CAAGACATTGAGACATAGAGTCCAAATTGGAGATGTGGGGACATAGTTTACAGAGAATC 780
 Qy 261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
 Db 781 CACTGCCGAGTGAACAAGTAAAGGGGTTTACTGTTTACAGTACGACGACCAGAGATA 840
 Qy 281 ValSerGlyLeuArgAspAsnThrIleLysIleTyrAspLysAsnThrLeuGluCysLys 300
 Db 841 GTACGGCGCTTCGACACACACCATCAAGATCTGGATAAAGACACACTGGAAATGCAAG 900
 Qy 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
 Db 901 CGGATTTCTACGGGGCCACACGGGCTCCGTCCTGTCTGCAGTACGATGAGAGGGTGATC 960
 Qy 321 IleThrGlySerSerAspSerThrValArgValTyrAspValAsnThrGlyGluMetLeu 340
 Db 961 ATCAGAGCTCTCTCAGACTCCACCGTCAGAGTGTGGGTGTAATGACGTGAGATGCTA 1020
 Qy 341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
 Db 1021 AACACATTGATTCACCACTGTGAAGCGGTTCTGCACCTCGCTTCAATTAATGSCATGATG 1080
 Qy 361 ValThrCysSerLysAspArgSerIleAlaValTyrAspMetAlaSerProThrAspIle 380
 Db 1081 GTACCTGTTTCCAAAGACCGCTTCCATCGCTGTGTGGGATATGGCTTCCCCAACTGACATC 1140
 Qy 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
 Db 1141 ACCCTCAGAGGGGTGTGTGTGGACACCGAGCTGCGGTCAATGTTGTAGACTTTGATGAC 1200
 Qy 401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTyrAsnThrSerThrCys 420
 Db 1201 AAGTACATCGTTTCTCGCTCTGGAGATAGAACCATAAAGGTGTGGAACACAAAGTACCTGT 1260
 Qy 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
 Db 1261 GAATTCGTAGAGACCTTAATGGCAACAGCGTGGCATCGCTGTTTGCAGTACAGAGAC 1320
 Qy 441 ArgLeuValValSerGlySerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460

```

1321 AGCGTGTGTGAGCGGCTCTCTGACACACCATCAGGCTGGGACATAGTGTGGA 1380
461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
1381 GCATGCTGCGAGTGTGGAGGCCATGAGGAGTGGTACGCTGCATTGGATTGATTAAC 1440
481 LysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTyrAspLeuValAla 500
1441 AAAAGGATAGTAGCGGAGCTATGATGGGAAATTAAGTGTGGGATCTTATGGCTGCT 1500
501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
1501 TTGACCCCGCTGCTCAGCAGGAGATCTCTGTCTCGGACACTTGTGGAGCATTTCTGGA 1560
521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
1561 AGAGTTTTCGGCTCCAGTTTGATGAATCCAGATTGTCTAGTAGTTTACATGATGACACA 1620
541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProArgSerPro 560
1621 ATTCTCATCTGGGACTTCCGAAATGATCCAGCTGCTCACGCTGAACCGCCGCTCCCT 1680
561 SerArgThrTyrThrTyrIleSerArg 569
1681 TCTCGGACATACACTTACATCTCCAGA 1707

```

13

4

293714 standard; DNA; 2175 BP.

293714;

-AUG-2000 (first entry)

cox protein FWDip coding sequence.

ubiquitin ligase; SCF; F-box protein; targeted degradation;
 stabilisation; proteolysis; drug discovery; gene therapy; cancer;
 oncoprotein; Huntington's disease; gene knockout; delivery systems;
 use; SS.

; musculus.

Location/Qualifiers

91..1800

/*tag= a

/product= F-box protein FWDip

00022110-A2.

APR-2000.

OCT-1999; 99WO-US23705.

OCT-1998; 98US-0103787.

RD | HARVARD COLLEGE.

u P, Howley P;

; 2000-317970/27.

SDB; AAY83254.

getting degradation of polypeptide useful for treating cancer and
 er proliferative disorders, involves conjugating polypeptide with
 quitin protein ligase or inhibiting ubiquitination using organic
 bound

im 10; Page 184; 18spp; English.

F-box proteins are a family of ubiquitin ligases (SCF ubiquitin
 ases) which can be used for the targeted degradation of a target

CC polypeptide in vivo. Targetted degradation is achieved by expressing
 CC the ubiquitin ligase in a cell linked to the interaction domain of
 CC the target polypeptide and thereby recruiting the target polypeptide
 CC to the ubiquitin ligase. Such methods are useful for decreasing or
 CC increasing the level of a target polypeptide and for creating and
 CC expressing a destabilized polypeptide which is subjected to SCF
 CC mediated proteolysis. Degrading any desired protein in a cell is
 CC useful for preventing or treating diseases caused by the presence of
 CC abnormal amount of the specific polypeptides, for drug discovery and
 CC for gene therapy. Diseases treated include cancer, by degradation of
 CC oncoproteins, Huntington's disease, other proliferative disorders and
 CC microbial infections. The method provides a quick and easy
 CC alternative to gene knockout technology. The target polypeptide can
 CC be degraded at all stages, or a specific stage, of development in the
 CC mature animal.

XX
 SQ Sequence 2175 BP; 589 A; 528 C; 567 G; 491 T; 0 other;

Alignment Scores:

Pred. No.: 3,83e-299 Length: 2175
 Score: 2997.00 Matches: 561
 Percent Similarity: 99.12% Conservative: 3
 Best Local Similarity: 98.59% Mismatches: 5
 Query Match: 98.78% Indels: 0
 DB: 21 Gaps: 0

US-09-601-168B-2 (1-569) x AA293714 (1-2175)

QY 1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
 Db 91 ATGGACCCGCGAGCGGCTGCTCAGGAGAAAGCGCTTAAGTTTATGAATTCCTCAGAG 150
 QY 21 ArgGluAspCysAsnAsnGlyGluProArgLysIleIleProGluLysAsnSerLeu 40
 Db 151 AGAGAAGACTGTAATAATGGCGAACCCCTAGAGAGATAATACCAGAGAAAGAAATTCATT 210
 QY 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
 Db 211 AGACAGACTTACAAACAGCTGTGCCAGGCTTTGCATATAACCAAGACAGATGATGCTAACAA 270
 QY 61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
 Db 271 AGCACTGCTATGAAGACTGAAATTTGTGTGGCCAAAGCCAAACTTGCCTAATGGCACTTCC 330
 QY 81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
 Db 331 AGCATGATTTGTGCCAAGCAGCGGAAACTCTCAGCAAGCTATGAGAAAGAAAGAGGCTG 390
 QY 101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
 Db 391 TGTGTCAAGTATTTTGGAGCAGTGGTCAGAGTCTGATCAAGTGGAAATTTGTAGAACACCTT 450
 QY 121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
 Db 451 ATATCCAAATGTGTCTACCTACAGCATGGGCGACATCAACTCTCTACTAAACCTATGCTG 510
 QY 141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
 Db 511 CAGAGGGATTTTCATNACTGCATCTGCCAGCAGCGGGTCTGGACCATCGCTGAGAACATT 570
 QY 161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr 180
 Db 571 CTGTCTACTTGGACGCCCAAGTCACTGTGTGTCTGTCTGAGCTCGTGTGCAAGGAATGCTGAC 630
 QY 181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
 Db 631 CGGCTGACGCTCGGACGGGCTGTGTGGAAAAGACTCATCGAGAGGATGCTCAGGACGGAC 690
 QY 201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220
 Db 691 TCTCTGTGGCAGGCTGCGCAGGCGCAGAGGCTGGGACAGTACTTATTCAAAAACAAA 750
 QY 221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240


```
|||||
751 CCTCTGATGAGAACGCTCTCCCACTCTTTTATAGAGCGTTTATCCTTAAATCAT 810
AC
XX
AAH89966;
XX
01-OCT-2001 (first entry)
XX
Human bone marrow cDNA, SEQ ID NO: 97.
DE
XX
Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
XX
antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
XX
immunosuppressive; gene therapy; cytokine cell proliferation;
XX
cell differentiation modulator; immune disorder; infection; cancer;
XX
human immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss.
OS
Homo sapiens.
XX
WO200153453-A2.
XX
26-JUL-2001.
XX
23-DEC-2000; 2000WO-US34960.
XX
21-JAN-2000; 2000US-0488725.
XX
25-APR-2000; 2000US-0552317.
XX
09-JUL-2000; 2000US-0598042.
XX
19-JUL-2000; 2000US-0620312.
XX
03-AUG-2000; 2000US-0653450.
XX
14-SEP-2000; 2000US-0662191.
XX
19-OCT-2000; 2000US-0693036.
XX
30-NOV-2000; 2000US-0250583.
XX
(HYSE-) HYSEQ INC.
XX
Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
XX
Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX
Zhao QA, Zhou P, Drmanac RT;
XX
WPI; 2001-488707/53.
XX
P-PSDB; AAM00847.
XX
Novel bone-marrow-expressed polynucleotides and polypeptides, useful
XX
for treating e.g. cancer and immune deficiency disorders -
XX
Claim 1; Page 274-275; 648pp; English.
XX
The present sequence is one of 251 novel human polynucleotides
XX
expressed in the bone marrow. The polynucleotide and the
XX
polypeptide encoded by it are useful in the treatment of various
XX
immune deficiencies and disorders. The deficiencies and disorders may
XX
be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
XX
infection, or may result from an autoimmune disorder; a coagulation
XX
disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
XX
suppression of an inflammatory response or treatment of a nervous
XX
system disorder such as Alzheimer's disease. Detection of the presence
XX
or increased expression of the polynucleotide or the protein it
XX
encodes is useful for the diagnosis and/or prognosis of one
XX
or more types of cancer. The polynucleotide and polypeptide can be
XX
used as nutritional sources or supplements and in the screening of
XX
chemical compounds as potential drugs.
XX
SQ Sequence 3622 BP; 961 A; 839 C; 874 G; 948 T; 0 other;
Alignment Scores:
Pred. No.: 2,12e-298 Length: 3622
Score: 2993.00 Matches: 569
Percent Similarity: 93.89% Conservative: 0
Best Local Similarity: 93.89% Mismatches: 0
Query Watch: 98.65% Indels: 37
DB: 22 Gaps: 1
US-09-601-168B-2 (1-569) x AAH89966 (1-3622)
Cy 1 MetAspProAlaGluAlaValLeuGlnGluLysAla-LeuLysPheMet----- 16
```

```

110 ATGGACCCGCGCGAGCGGTGCTGCAAGAGAGGACCTCAAGTTTATGTGCTCTATGCC 169
16 ----- 16
170 CAGGTCTCTGTGGCTGGGCTGCTCCAGCTGGCGGACAGCATGCCCTTCGCTGGCATGCCT 229
17 -----AenSerSerGluArgGluAspCy 24
230 GTATAACCCAGGCACTGGCGCACTCACAGCTTTCAGAAATTCCTCAGAGAGAAGACTG 289
24 sAsnAsnGlyGluProProArgLysIleIleProGluLysSerSerLeuArgGlnThrTy 44
290 TAATAATGCGAACCCCTTAGAAGATTAATACGAGAAGAAATTCACCTTAGACAGACATA 349
44 rAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAlaSerThrAlaMe 64
350 CAACAGCTGTGCCAGACTCTGCTTAAACCAAGAAACAGTATGTTTAGCAAGCACTGCTAT 409
64 tLysThrGluAenCysValAlaLysThrLysLeuAlaAsnGlyThrSerSerMetIleVa 84
410 GAAGACTGAGAAATGTGTGCGCAACAAACCTTCCCAATGGCACTTCCAGTATGATGT 469
84 lProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeuCysValLysTy 104
470 GCCCAAGCAACGGAACTCTCAGCAAGCTATGAAAGGAAGGAAGCACTGTGTCAATA 529
104 rPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeuIleSerGlnMe 124
530 CTTTGAGCAGTGTGTCAGAGTCAGATCAAGTGAATTTGTGGAACATCTTATATCCCAAT 589
124 tCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeuGlnArgAspPh 144
590 GTGTCTATACCAACATGGGCATTAACCTGATCTTAAACCTATGTTGGAGAGATTT 649
144 elleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAenIleLeuSerTyrLe 164
650 CATAACTGCTCTGCCAGCTCGGGATGGATCATATCGCTGAGAAATTCCTGTATACCT 709
164 uAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyrArgValThrSe 184
710 GGATGCCAAATCACTATGTGCTGCTGAACCTGTGCAAGGAATGGTACCGAGTGACCTC 769
184 rAspGlyMetLeuTrpLysLeuIleGluArgMetValArgThrAspSerLeuTrpAr 204
770 TGATGGCATGCTGTGAAGAGCTTATCAGAGAAATGGTCAGACAGATTCCTCTGGAG 829
204 gGlyLeuAlaGluArgGlyTrpGlyGlnTyrLeuPheLysAsnLysProProAspGl 224
830 AGGCCTGGCAGAACGAGAGGATGGGACAGATTTATTTCAAAAAACAAACCTCTCTGACGG 889
224 yAenAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleGl 244
890 GAATGCTCTCCCACTCTTTTATAGACACTTATCTCTAAATATACAGACATTTGA 949
244 uThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSe 264
950 GACAAATAGAAATCTAATTGGAGATGTGGAAGACATAGTTTACAGAGAATTCACGCGGAAG 1009
264 rGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLe 284
1010 TGAACAAGCAAGCAAGGAGTTACTGTTTACAGATGATGATGCAAAAAATAGTAAGCGGCT 1069
284 uArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLysArgIleLeuTh 304
1070 TCAGAGACAACAACATCAAGATCTGGATTAATAAACACATTTGGATGCAAGCAATCTCAC 1129
304 rGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIleIleThrGlySe 324
1130 AGGCCATACAGGTTCAAGTCTCTGCTCCAGTATGATGAGAGAGTGATCATAACAGGATC 1189
324 rSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIl 344

```

RESULT 15

```

AAK51717
ID AAK51717 standard; cDNA; 2366 BP.
XX AAK51717;
XX XX
XX 06-NOV-2001 (first entry)
XX Human polynucleotide SEQ ID NO 262.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorder; arthritis; inflammation; ss.
XX Homo sapiens.
XX
XX WO200157190-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US04098.
XX

```

FEB-2000; 2000US-0496914.
 APR-2000; 2000US-0560875.
 JUN-2000; 2000US-0598075.
 JUL-2000; 2000US-0620325.
 SEP-2000; 2000US-0654936.
 SEP-2000; 2000US-0654936.
 OCT-2000; 2000US-0663561.
 OCT-2000; 2000US-0693325.
 NOV-2000; 2000US-0728422.

(XSE-) HYSEQ INC.

g YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 e AU, Yang Y, Wehrman T, Goodrich R;

; 2001-476283/51.

SSDB; AAK78584.

leic acids encoding polypeptides with cytokine-like activities,
 ful in diagnosis and gene therapy -

im 1; Page 1183-1186; 6221pp; English.

invention relates to polynucleotides (AAK51456-AAK53435) and the
 oded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
 okine, cell proliferation or cell differentiation or which may induce
 uction of other cytokines in other cell populations. The
 ucleotides and polypeptides are useful in gene therapy, vaccines or
 ide therapy. The polypeptides have various cytokine-like activities,
 stem cell growth factor activity, haematopoiesis regulating
 ivity, tissue growth factor activity, immunomodulatory activity and
 ivin/inhibin activity and may be useful in the diagnosis and/or
 atment of cancer, leukaemia, nervous system disorders, arthritis and
 lamation.

e: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 MM80020) are omitted as the relevant pages from the sequence listing
 e missing at the time of publication.

quence 2366 BP; 661 A; 519 C; 594 G; 592 T; 0 other;

t Scores:

..	1.27e-298	Length:	2366
	2992.50	Matches:	569
Similarity:	90.03%	Conservative:	0
al Similarity:	90.03%	Mismatches:	0
atch:	98.63%	Indels:	63
	22	Gaps:	1

1-168B-2 (1-569) x AAK51717 (1-2366)

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLys----- 14
 127 ATGACCGCGCGAGCGGTCTGCAAGAGAAGGCACTCAAGTTTATGATGGAGTTTCGC 186
 14 ----- 14
 187 TCTTGGTGCCAGCGTGGATACAAATGGCGCGATCTCGGCTCACCGCAACCTCCACCTCC 246
 14 ----- 14
 247 CGGGTTCAATGCTCTATGCCAGGTCTCTGTGGCTGGGCTGCTCCAGCCTGGCGGACAGC 306
 15 -----PheMetAsn 17
 307 ATGCTTCCGTCGATGCTGTATATACCCAGGAGCTGGCGCACTCACAGCTTTTCATGAAT 366
 18 SerSerGluArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLys 37
 367 TCCTCAGAGAGAAGACTGTAATAATGGCGAAGCCCTAGAGAAGATAATACCAGAGAAG 426
 38 AsnSerLeuArgGlnThrThrAspSerCysAlaArgLeuCysLeuAsnGlnGluThrVal 57
 427 AATTCGCTTAGACAGACATACACAGCTGTCCAGACTCTCTGCTTTAAACCAAGAAACAGTA 486

QY 58 CysLeuAlaSerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsn 77
 |||||
 Db 487 TGTTTAGCAAGCACTGCTATGAAGACTGAGAAATCTGTGGCCAAACAAACCTTCCCAAT 546
 |||||
 QY 78 GlyThrSerSerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGlu 97
 |||||
 Db 547 GGCACCTTCCAGTATGATTGTGCCAAGCAACGAAACTCTCAGCAAGCTATGAAAGGAA 606
 |||||
 QY 98 LysGluLeuCysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheVal 117
 |||||
 Db 607 AAGGAACCTGTGTGTCRAATACCTTTGAGCAGTGCTCAGAGTCAGATCAAGTGAATTTGTG 666
 |||||
 QY 118 GluHisLeuIleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLys 137
 |||||
 Db 667 GAACATCTTATATCCAAATGTGTCTATACCAACATGGGCACATAAACTCGTATCTTAAA 726
 |||||
 QY 138 ProMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAla 157
 |||||
 Db 727 CCTATTGTGCAGAGATTTCTATAACTGCTCTGCCAGCTCGGGGATTGGATCATATTGCT 786
 |||||
 QY 158 GluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLys 177
 |||||
 Db 787 GAGAACATTTCTGTCTATACCTGGATGCCAAATCACTATGTGCTGCTGAACTTTGTGCAAG 846
 |||||
 QY 178 GluTrpTyrArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetVal 197
 |||||
 Db 847 GAATGGTACCGAGTGACCTCTGATGCGATGCTGTGGAAGAGCTTATCGAGAGAATGGTC 906
 |||||
 QY 198 ArgThrAspSerLeuTrpArgGlyLeuAlaGluArgGlyTrpGlyGlnTyrLeuPhe 217
 |||||
 Db 907 AGCAGAGATTTCTCTGTGAGAGCGCTGCGAGAACGAGAGATGGGACAGTATTATTTC 966
 |||||
 QY 218 LysAsnLysProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrPro 237
 |||||
 Db 967 AAAACAAACCTCTCTGACGGGAATGCTCTCCCACTCTTTTATAGAGCATTATCTCT 1026
 |||||
 QY 238 LysIleIleGlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeu 257
 |||||
 Db 1027 AAAATTATACAAGACATTGAGACAATAGAACTTAATTGGAGATGTGGAAGACATAGTTTA 1086
 |||||
 QY 258 GlnArgIleHisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAsp 277
 |||||
 Db 1087 CAGAGAAATTCATGCGGAAGTGAACAAAGCAAGAGGAGTTTACTGTTTACAGTATGATGAT 1146
 |||||
 QY 278 GlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeu 297
 |||||
 Db 1147 CAGAAAATAGTAAGCGGCTTCGAGACAAACACAAATCAAGATCTGGGATAAAACACATTG 1206
 |||||
 QY 298 GluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGlu 317
 |||||
 Db 1207 GAATGCAAGCGAAATTCACAGGCCATACAGGTTTCAGTCTCTGTCTCCAGTATGATGAG 1266
 |||||
 QY 318 ArgValIleIleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGly 337
 |||||
 Db 1267 AGAGTGATCATAAACAGGATCATCGGATTCACGGTCAGAGTGTGGGATGTAATACAGGT 1326
 |||||
 QY 338 GluMetLeuAsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsn 357
 |||||
 Db 1327 GAAATGCTAAACACGTTGATTCAACCATTTGTGAAGCAGTCTGCACCTTCGTTTCAATAAT 1386
 |||||
 QY 358 GlyMetMetValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerPro 377
 |||||
 Db 1387 GGCATGATGGTGACCTGCTCCAAAGATCGTTCATTTGCTGTATGGGATATGGCCTCCCA 1446
 |||||
 QY 378 ThrAspIleThrLeuArgValLeuValGlyHisArgAlaAlaValAsnValValAsp 397
 |||||
 Db 1447 ACTGCATTAACCTCCGAGGGTGTGTGTCGACACCGAGCTGCTGTCAATGTTGTAGAC 1506
 |||||
 QY 398 PheAspAspLysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThr 417
 |||||
 Db 1507 TTTGATGACAGTACATTTGTTCTGCTCTGGGGATAGAACTATTAAGGTATGGAAACACA 1566
 |||||

.
418 SerThrCysGluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGln 437
|||||
1567 AGTACTTGTGAATTCTGAAGGACCTTAAATGGACACAAACGAGGCATTGCCTGTTTGCG 1626
438 TyrArgAspArgLeuValSerGlySerSerAspAsnThrIleArgLeuTrpAspIle 457
|||||
1627 TACAGGACAGGCTGGTAGTGAGTGGCTCATCTGACCAACACTATCAGATTATGGGACATA 1686
458 GluCysGlyAlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArg 477
|||||
1687 GAATGTGGTGCATGTTTACGAGTGTAGAAAGGCCATGAGCAATTGGTGGTTGTATTGGA 1746
478 PheAspAsnLysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeu 497
|||||
1747 TTGTATACAAAGAGGATAGTCACTGGGGCTATGATGGAAAAATTAAAGTGGGATCTT 1806
498 ValAlaAlaLeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGlu 517
|||||
1807 GTGGCTGCTTTGGACCCCGCTGCTCTGCAGGGACACTCTGTCTACGGACCCCTTGTGGAG 1866
518 HisSerGlyArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHis 537
|||||
1867 CATTCGGGAAGAGTTTTTCGACTACAGTTTGATGAATTCAGATTGTCCAGATTGTCCACAT 1926
538 AspAspThrIleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProPro 557
|||||
1927 GATGACACAAATCCTCATCTGGGACTTCTTAATGATCCAGCTGCCCAAGCTGAACCCCC 1986
558 ArgSerProSerArgThrTyrThrIleSerArg 569
|||||
1987 CGTTCCCTTCTCGAACATACACCTACATCTCCAGA 2022

completed: October 22, 2003, 15:03:20

s : 422 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

main - nucleic search, using frame_plus_p2n model

October 22, 2003, 16:26:53 ; Search time 395 Seconds

(without alignments)
3963.493 Million cell updates/sec

US-09-601-168b-2

score: 3034
= 1 MDPRAVLQKALKEFNSE.....PAAQAEPRSPSRITYVISR 569

table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

d: 1792395 seqs, 1340900451 residues

umber of hits satisfying chosen parameters: 3584790

DB seq length: 0

DB seq length: 2000000000

rocessing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

line parameters:

frame_p2n.model -DEV-xmlp
?_1/USFTO.spool_p/US09601168/runat_22102003.100548.17891/app_query.fasta.1.711
ished Applications NA -QRMFIstap -SUFFIX=rnpb -MINMATCH=0.1
-0 -LOOPEXT=0 -UNITFS=bits -START=1 -END=1 -MATRIX=BLOSUM62
uman40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-0 -ALIGN=15 -MODS=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
2000000000 -USER=US09601168 @CGN.1.1.95 @runat_22102003.100548.17891
-ICPU=3 -NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
3 -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Published Applications NA:

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTVS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

ed. No. is the number of results predicted by chance to have a
ore greater than or equal to the score of the result being printed,
id is derived by analysis of the total score distribution.

SUMMARIES

Query	Score	Match	Length	ID	Description
3034	100.0	1710	14	US-10-038-010-7	Sequence 7, Appli

2	3034	100.0	2151	12	US-09-832-161-17	Sequence 17, Appli
3	3034	100.0	2151	13	US-10-042-417-1	Sequence 1, Appli
4	2384.5	78.6	4230	12	US-09-832-161-15	Sequence 15, Appli
5	1163	38.3	657	14	US-10-023-530-1	Sequence 1, Appli
6	1155.5	38.1	951	9	US-09-764-848-15	Sequence 15, Appli
7	1155.5	38.1	951	12	US-10-222-020-15	Sequence 15, Appli
8	1155.5	38.1	951	14	US-10-116-016-15	Sequence 15, Appli
9	689	22.7	479	9	US-09-864-761-2778	Sequence 2778, Ap
10	644.5	21.2	1881	10	US-09-213-888-20	Sequence 20, Appli
11	644.5	21.2	1881	10	US-09-328-877A-13	Sequence 20, Appli
12	640	21.1	1620	12	US-10-245-618-20	Sequence 13, Appli
13	640	21.1	1659	12	US-10-245-618-7	Sequence 7, Appli
14	640	21.1	1770	12	US-10-245-618-17	Sequence 17, Appli
15	640	21.1	1884	12	US-10-245-618-5	Sequence 5, Appli
16	640	21.1	2001	10	US-09-213-888-26	Sequence 26, Appli
17	640	21.1	2001	10	US-09-328-877A-26	Sequence 26, Appli
18	640	21.1	2010	10	US-09-213-888-24	Sequence 24, Appli
19	640	21.1	2010	10	US-09-328-877A-24	Sequence 24, Appli
20	640	21.1	2063	12	US-10-245-618-34	Sequence 34, Appli
21	640	21.1	2124	12	US-10-245-618-1	Sequence 1, Appli
22	640	21.1	2255	12	US-10-245-618-35	Sequence 35, Appli
23	640	21.1	3550	10	US-09-213-888-1	Sequence 1, Appli
24	640	21.1	3571	10	US-09-328-877A-1	Sequence 1, Appli
25	640	21.1	3571	10	US-09-213-888-2	Sequence 2, Appli
26	640	21.1	3571	10	US-09-328-877A-2	Sequence 2, Appli
27	575	19.0	2037	10	US-09-801-368-313	Sequence 313, App
28	552.5	18.2	2091	14	US-10-128-714-2506	Sequence 2506, Ap
29	552.5	18.2	2091	14	US-10-128-714-7506	Sequence 7506, Ap
30	552.5	18.2	2167	14	US-10-128-714-1506	Sequence 1506, Ap
31	552.5	18.2	2167	14	US-10-128-714-6506	Sequence 6506, Ap
32	552.5	18.2	4167	14	US-10-128-714-506	Sequence 506, App
33	552.5	18.2	4167	14	US-10-128-714-5506	Sequence 5506, App
34	519.5	17.1	2208	9	US-10-032-585-6658	Sequence 6658, Ap
35	517	17.0	370	9	US-09-864-761-19492	Sequence 19492, A
36	414	13.6	524	12	US-10-029-386-4953	Sequence 4953, Ap
37	405	13.3	350	12	US-10-029-386-18709	Sequence 18709, A
38	392	12.9	2237	10	US-09-994-485-7	Sequence 7, Appli
39	392	12.9	2237	10	US-09-832-232-11	Sequence 11, Appli
40	344	11.3	1548	14	US-10-128-714-7213	Sequence 7213, Ap
41	342	11.3	319	10	US-09-960-352-2934	Sequence 2934, Ap
42	338.5	11.2	1923	14	US-10-102-806-190	Sequence 190, App
43	338.5	11.2	3747	14	US-10-197-666A-143	Sequence 143, App
44	338.5	11.2	3864	14	US-10-197-666A-149	Sequence 149, App
45	337	11.1	3465	10	US-09-994-485-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-10-038-010-7
; Sequence 7, Application US/10038010
; Publication No. US20030040089A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, Legrain
; TITLE OF INVENTION: Protein-protein interactions in adipocyte cells
; FILE REFERENCE: B4767A
; CURRENT APPLICATION NUMBER: US/10/038,010
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 60/259,377
; PRIOR FILING DATE: 2002-07-23
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 7
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(1710)
; OTHER INFORMATION: Human beta-TrCP1 : F-box containing protein with 7 WD40 repeats;
; OTHER INFORMATION: Part of SCF (Skp1/Cullin/F-box) complex E3 ubiquitin ligase; Impl
; OTHER INFORMATION: located in the degradation of b-catenin and IkBa

38-010-7

nt Scores:

0 Length: 1710
3034.00 Matches: 569
Similarity: 100.00% Conservativeness: 0
cal Similarity: 100.00% Mismatches: 0
atch: 100.00% Indels: 0
14 Gaps: 0

31-168B-2 (1-569) x US-10-038-010-7 (1-1710)

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
1 ATGGACCCGGCGAGCGGTGCTGCAAGAGAGAGGACCTCAAGTTATGAATTCCTCAGAG 60
21 ArgGluAspCysAsnAsnGlyGluProProArgLysIleProGluLysAsnSerLeu 40
61 AGAGAAAGACTGTAATAATGCGAAGCCCTAGGAAGATAATACCAAGAGAAGATTCACCT 120
41 ArgGlnThrTyrrAsnSerCysAlaArgLeuLysLeuAsnGlnGluThrValCysLeuAla 60
121 AGACAGACATACACAGCTGTCCAGACTCTGCTTTAAACCAAGAAACAGTATGTTTAGCA 180
61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
181 AGCACTGCTATGAAGACTGAGAATTTGTGTGGCCAAAACAAACATTCGCAATGGCACTTC 240
81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrrGluLysGluLysGluLeu 100
241 AGTATGATTGTGCCCAAGCAAGCAAGCAACTCTCAGCAGCTATGAAAGAAAGAAAGAACTG 300
101 CysValLysTyrrPheGluGlnThrSerGluSerAspGlnValGluPheValGluHisLeu 120
301 TGTGTCAATACCTTTGAGCAGTGGTTCAGAGTCCAGATCAAGTGAATTTGTGGAACATCTT 360
121 IleSerGlnMetCysHisTyrrGlnHisGlyHisIleAsnSerTyrrLeuLysProMetLeu 140
361 ATATCCCAATGTGTCATACCAATGAGCAATGGCAATGAACTCGTATCTTAAACCTATGTTG 420
141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
421 CAGAGAGATTCATACCTCTCTGCCAGCTCGGGATTTGGATCATATCGCTGAGAACATTT 480
161 LeuSerTyrrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGlnTyrr 180
481 CTGTCTATCTGGATGCCAAATCACTATGCTGTCTGAACCTTGTGTGCAAGGAATGTGTAC 540
181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
541 CGAGTGACCTCTGATGGCATGCTGTGGAAAGAGCTTATCGAGAGATGCTCAGGACAGAT 600
201 SerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrrLeuPheLysAsnLys 220
601 TCTCTGTGAGAGAGCGCTGCGAAGCAAGAGAGATGGGACAGTATTTATTCAAAACAAA 660
221 ProProAspGlyAsnAlaProProAsnSerPheTyrrArgAlaLeuTyrrProLysIleIle 240
661 CCTCTCTGAGCGGAATGCTCTCTCCCACTCTTTTATAGGACATTTATCTTAAATTTATA 720
241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
721 CAAGACATTTGAGACATAGAACTAATTTGAGATGTGGAAGACATAGTTTACAGAGAATT 780
261 HisCysArgSerGluThrSerLysGlyValTyrrCysLeuGlnTyrrAspAspGlnLysIle 280
781 CACTGCGGAAGTGAACAAAGAGAGATTTACTGTTTACAGTATGATGATCAGAAAAATA 840
281 ValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLys 300
841 GTAAGCGGCTTCGACACACAACTCAGATCTGGGATAAAAACACATTTGAATGCAAG 900
301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrrAspGluArgValIle 320

Db 901 CGAATTCCTCAGAGGCCATACAGGTTCAAGTCTCTGTCTCCAGTATGATGAGAGAGTGATC 960
QY 321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
Db 961 ATAACAGAGTATCGGATTCACCGGTGAGAGTGGGATGTAATACAGGTGAATGCTA 1020
QY 341 AsnThrLeuIleHisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
Db 1021 AACAGTTGATTCACCATTTGAAGCAGTCTGCACCTGGCTTCAATATATGCGCATGATG 1080
QY 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
Db 1081 GTGACCTGCTCCAAAGATCGTTCATTTGCTATGAGGATATGGCTCCCACTGACATT 1140
QY 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
Db 1141 ACCCTCCGAGGGTCTGGTCCGACACCGAGCTGCTCAATGTTGTAGACTTTTGTAGAC 1200
QY 401 LysTyrrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
Db 1201 AAGTACATTTGTTCTGCTCTGGGATAGAACTATTAAGGTATGGAACACAGTACTTGT 1260
QY 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrrArgAsp 440
Db 1261 GAATTTGTAAGGACCTTAAATGGACACAAACGAGGCATTTGCTGTTTGCAGTACAGGAC 1320
QY 441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
Db 1321 AGGCTGTGTAGTGGTGGCTCATCTGCACACACTATCAGATTATGGGACATAGATGTGGT 1380
QY 461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
Db 1381 GCATGTTTACAGAGTGTAGAAAGGCCATGAGGAATTTGGTGGCTTGTATTTCGATTGATAAC 1440
QY 481 LysArgIleValSerGlyValTyrrAspGlyLysLysValTrpAspLeuValAlaAla 500
Db 1441 AAGAGATAGTCAAGTGGGGCTTATGATGGAATAATTAAGTGGGATCTTTGTGGCTGCT 1500
QY 501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
Db 1501 TTGGACCCCGTGTCTCTGAGGAGACACTGTCTACGGACCTTGTGGAGCATTCGGGA 1560
QY 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
Db 1561 AGAGTTTTCGACTACAGTTTGATGAATTCAGATTGTCAGTAGTTCACATGATGACACA 1620
QY 541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
Db 1621 ATCTCATCTGGGACTTCTTAAATGATCCAGCTGCCAAGCTGAACCCCGCTTCCCT 1680
QY 561 SerArgThrTyrrThrTyrrIleSerArg 569
Db 1681 TCTCGAACATACACCTACATCTCCAGA 1707

RESULT 2

US-09-832-161-17
; Sequence 17, Application US/09832161
; Publication No. US20030166587A1
; GENERAL INFORMATION:
; APPLICANT: Manning, Anthony M.
; APPLICANT: Mercurio, Frank
; APPLICANT: Amit, Sharon
; APPLICANT: Ben-Neriah, Yinon
; APPLICANT: Davis, Matti
; APPLICANT: Hatzuba, Ada
; APPLICANT: Lavon, Iris
; APPLICANT: Yaron, Avraham
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING ACTIVATION OF
; FILE REFERENCE: 860098.427
; CURRENT APPLICATION NUMBER: US/09/832,161
; CURRENT FILING DATE: 2001-04-09

APPLICATION NUMBER: 09/210,060

FILING DATE: 1998-12-10

R OF SEQ ID NOS: 30

ARE: PatentIn Ver. 2.0

NO 17

TH: 2151

: DNA

NISM: Homo sapiens

2-161-17

t Scores:

.: 0 Length: 2151

Matches: 569

Similarity: 3034.00

Conservative: 0

Mismatch: 100.00%

Indels: 0

Gaps: 0

1-168B-2 (1-569) x US-09-832-161-17 (1-2151)

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20

70 ATGACCCGGCGGAGGCGTCTGCAAGAGAGCAGCTCAAGTTTATGAATTCCTCAGAG 129

21 ArgGluAspCysAsnAsnGlyGluProProArgLysIleProGluLysAsnSerLeu 40

130 AGAAGAGACTGTAATAATGGCGAACCCCTAGGAAGATAATACCAGAGAAGAAATTCATT 189

41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60

190 AGACAGACATACACAGCTGTCGCTTAAACCAAGAAACAGATATGTTTAGCA 249

61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80

250 AGCACTGCTATGAGACTGAGANTGTGGCCAAACAAACTTGGCAATGGCACTTCC 309

81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100

310 AGTATGATTGTGCCCAAGCAAACTCTCAGCAAGCTATGAAAGGAAAGGAACTG 369

101 CysValLysTyrPheGluGlnThrSerGluSerAspGlnValGluPheValGluHisLeu 120

370 TGTCTCAATACTTTGAGCAGTGTGACAGTCAGATCAAGTGGAAATTTGTGGAACTCTT 429

121 IleSerGlnMetCysHisTyrGlnHisLysIleAsnSerTyrLeuLysProMetLeu 140

430 ATATCCCAATGCTCATTACCAACATGGGCACATAAATCTGATCTTTAAACCTATGTTG 489

141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160

490 CAGAGAGATTTCATAACTGCTCTGCCAGCTCGGGGATTGGATCATATCGCTGAGAACATT 549

161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluThrTyr 180

550 CTGTCACTACCTGGATGCCAATCACTATGTCTCTGAACTTGTGTCGAAGGAAATGGTAC 609

181 ArgValThrSerAspGlyMetLeuTriLysLysLeuIleGluArgMetValArgThrAsp 200

610 CGAGTGACCTCTGATGGCATGCTGTGGAAGAGCTTTCGAGAGAAATGGTCAGACAGAT 669

201 SerLeuTrpArgGlyLeuAlaGluArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220

670 TCTCTGTGGAGAGCGCTGGCAGAACGAGAGATGGGAGATGTTATTCAAAACAAA 729

221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240

730 CCTCCTGACGGGAATGCTCTCCCAACTCTTTTATAGACACATTTATCTCTAAATATATA 789

241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260

790 CAAACATTTGACACATAGAAATCAATTTGGAGATGTGGAACACATATGTTACAGAGAA 849

261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280

Db 850 CACTGCCGAGTGAACAAGCAAGAGTTTACTGTTTACAGTATGATGATCAGAAATA 909

QY 281 ValSerGlyLeuArgAspAsnThrIleLysIleTyrAspLysAsnThrLeuGluCysLys 300

Db 910 GTAAGCGGCTTCGAGACAACAACATCAAGATCTGGGATAAAACACACATTGGAATGCAAG 969

QY 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320

Db 970 CGAATTTCTCAGAGGCCATACAGTTTCAGTCTCTGTTCCAGTATGATGAGAGATGATC 1029

QY 321 IleThrGlySerSerAspSerThrValArgValTyrAspValAsnThrGlyGluMetLeu 340

Db 1030 ATAACAGGATCATCGGATTCACGCTCAGAGTGGGATGTAATACAGTGAATGCTA 1089

QY 341 AsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360

Db 1090 AACAGTTCGATTCCACATTGTGAAGCAGTTCTGCACCTTCGTTTCAATAATGSCATGATG 1149

QY 361 ValThrCysSerLysAspArgSerIleAlaValTyrAspMetAlaSerProThrAspIle 380

Db 1150 GTGACCTGCTCCAAAGATCGTTCCATTGCTGTATGGGATATGGCTCCCAACTGCATT 1209

QY 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400

Db 1210 ACCCTCCGAGGGTGTGGTGGACACCGAGCTGCTCAATGTTGTAGACTTTGATGAC 1269

QY 401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTyrAsnThrSerThrCys 420

Db 1270 AAGTACATTGTTTCTGCATCTCGGGATAGAACTATAAAGGTATGGAACACAAGTACTTGT 1329

QY 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440

Db 1330 GAATTTGTAAAGACCTTAAATGGACACAAACGAGGCAATTCCTGTTTGCAGTACAGGAC 1389

QY 441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTyrAspIleGluCysGly 460

Db 1390 AGGCTGGTAGTCAGTGGCTCACTGCACAACACTATCAGATTATGGACATAGAAATGTGCT 1449

QY 461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480

Db 1450 GATGTTTACAGAGTGTAGAGGCCATGAGGAATTTGGTSCGTGTATTTCGATTTTGATAAC 1509

QY 481 LysArgIleValSerGlyValTyrAspGlyLysIleLysValTyrAspLeuValAlaAla 500

Db 1510 AACAGGATAGTCAGTGGGSCCTATGATGGAAAAAATTAAGTGTGGATCTTGTGCTGCT 1569

QY 501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520

Db 1570 TTGGACCCCGTGTCTCGAGGACACTCTGCTACGGACCTTGTGGAGCATTCGGA 1629

QY 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerSerHisAspAspThr 540

Db 1630 AGAGTTTTCGACTACAGTTTCATGAATTCAGATTTCAGTAGTTTCATATGATGACACA 1689

QY 541 IleLeuIleTyrAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560

Db 1690 ATCCTCATCTGGACATCTCTAAATGATCCAGCTGCCAGCTGAACCCCCCTTCCCT 1749

QY 561 SerArgThrTyrThrTyrIleSerArg 569

Db 1750 TCTCGAACATACACCTACATCTCCAGA 1776

RESULT 3

US-10-042-417-1

; Sequence 1, Application US/10042417

; Publication No. US20020123082A1

; GENERAL INFORMATION:

; APPLICANT: Pagano, M.

; TITLE OF INVENTION: METHODS TO IDENTIFY COMPOUNDS USEFUL FOR THE TREATMENT OF

; Proliferative and Differentiative Disorders

; FILE REFERENCE: 5914-090-999

; CURRENT APPLICATION NUMBER: US/10/042,417

NT FILING DATE: 2002-01-07
APPLICATION NUMBER: 60/260,179

FILING DATE: 2001-01-5

R OF SEQ ID NOS: 89

ARE: PatentIn Ver. 2.0

NO 1

TH: 2151

: DNA

UNISM: Homo sapiens

2-417-1

t Scores:

Length: 2151
Matches: 569
Similarity: 3034.00
Conservative: 0
Mismatch: 100.00%
Indels: 0
Gaps: 0

1-168B-2 (1-569) x US-10-042-417-1 (1-2151)

1 MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu 20
70 ATGGACCGGCGGAGGGGGTCTGCAAGAGAGAGGCACTCAAGTTTATGAATTCCTCAGAG 129
21 ArgGluAspCysAsnAsnAnglyGluProProArgLysIleIleProGluLysAsnSerLeu 40
130 AGAAGAGACTGTAAATAATGGCGAACCCCTAGGAAGATAATACAGAGAGAGAAATTCACTT 189
41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla 60
190 AGACAGACATACACAGCTGTGCAGACTCTGCTTAAACCAAGAACAGATATGTTAGCA 249
61 SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer 80
250 AGCACTGCTATGAAGACTGAGAAATTGTGTGGCCAAACAAACTTGCATATGCCACTTCC 309
81 SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeu 100
310 AGTATGATTGGCCCAAGCAACGGAACTCTCAGCAAGCTATGAAAGGAAAGGAACCTG 369
101 CysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGluHisLeu 120
370 TGTGTCAAAATCTTTGAGCAGTGTGAGAGTCAGATCAGTGGNAATTGTGGACATCTT 429
121 IleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysProMetLeu 140
430 ATATCCCAAATGTGTCATTACCAACATGGGCACATAAACTCGTATCTTAAACCTATGTTG 489
141 GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle 160
490 CAGAGAGATTTCATACTGCTCTCCAGCTCGGGGATTTGATCATATCGCTGAGAACATT 549
161 LeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTyrTyr 180
550 CTGTCACTACCTGGATGCCAATCACTATGTGCTGCTGAACCTGTGTGCAAGGAATGCTAC 609
181 ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp 200
610 CGAGTGACCTCTGATGGCATGCTGTGGAAGAAGCTTATCGAGAGAAATGGTCCAGGACAGAT 669
201 SerLeuTrpArgGlyLeuAlaGluArgGlyTrpGlyGlnTyrLeuPheLysAsnLys 220
670 TCTCTGTGGAGAGCCCTGGCAGACAGAGAGATGGGACAGATATTATTATCAAAAAACAA 729
221 ProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIle 240
730 CCTCTGACGGGAATGTCTCTCCCACTCTTTTATAGACACTTTATCTTAAATTTATA 799
241 GlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIle 260
790 CAAGACATTGAGACAATAGAAATCTAATTGGAGATGTGGAACATAGTATTACAGAGAAAT 849

QY 261 HisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIle 280
Db 850 CACTCCCGAAGTGAACAAGCAAGAGAGTCTTACTGTTTACAGATATGATCATCAGAAAAATA 909
QY 281 ValSerGlyLeuArgAspAsnThrIleLysIleIleIleLysAsnThrLeuGluCysLys 300
Db 910 GTAAGCGGCTTCGAGACAACACATCAAGATCTGGGATAAANAACACATTGGAAATGCAAG 969
QY 301 ArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArgValIle 320
Db 970 CGAATTCTCAGAGGCATACAGGTCAGTCTCTCTCCAGTATGATGAGAGAGTGATC 1029
QY 321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
Db 1030 ATAACAGGATCATCGGATTCACGCTCAGAGTGGGATGTAATAACAGGTGAAATGCTA 1089
QY 341 AsnThrLeuIleHisHisCysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMet 360
Db 1090 AACACGTTGATTCAACATTGTGAAGCAGTCTCGCACTTGCCTTCAATAATGGCATGATG 1149
QY 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
Db 1150 GTGACCTGCTCAAAAGATCGTTCCATTGCTGTATGGGATATGGCCTCCCAACTGACATT 1209
QY 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPheAspAsp 400
Db 1210 ACCCTCCGAGGGTCTGCTGGACACCGAGCTGCTGTCATGTTGTAGACTTTTGATGAC 1269
QY 401 LysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThrCys 420
Db 1270 AAGTACATTGTTCTGCACTCTGGGATAGAACTATAAAGGTATGAAACACAAGTACTTGT 1329
QY 421 GluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAsp 440
Db 1330 GAATTGTAAGGACCTTAAATGGACACAAACGAGGCACTTGCTGTTCAGTACAGGAC 1389
QY 441 ArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGly 460
Db 1390 AGCTGGTAGTGAGTGGCTCATCTGACAACTATCAGATTATGGACATAGAAATGTGT 1449
QY 461 AlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAspAsn 480
Db 1450 GCATGTTTACGAGTGTAGAAAGGCCATCAGGAATTGGTGGCTGTGATTCGATTTCGATAAC 1509
QY 481 LysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAla 500
Db 1510 AAGAGGATAGTCAGTGGGGCCTATGATGGAATAATTAAGTGTGGGATCTTGTGGCTGCT 1569
QY 501 LeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGly 520
Db 1570 TTGGACCCCGTGCTCTCGAGGGACACTCTGTCTACGGACCTTGTGGAGCATTCGGGA 1629
QY 521 ArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThr 540
Db 1630 AGAGTTTTTCGACTACAGTTTGATGAATTCAGATTGTTCAGTAGTTCATGATGACACA 1689
QY 541 IleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProProArgSerPro 560
Db 1690 ATCTCTCATCTGGGACTTCTTAATGATCCAGTCCCAAGCTGAACCCCGGTTCCGCT 1749
QY 561 SerArgThrTyrThrTyrIleSerArg 569
Db 1750 TCTCGAACATACACCTACATCTCCAGA 1776

RESULT 4

US-09-832-161-15

; Sequence 15, Application US/09832161

; Publication No. US20030166587A1

; GENERAL INFORMATION:

; APPLICANT: Manning, Anthony M.

; APPLICANT: Mercurio, Frank

; APPLICANT: Amit, Sharon

; APPLICANT: Ben-Neriah, Yinon

APPLICANT: Davis, Matti
 APPLICANT: Hatzubai, Ada
 APPLICANT: Navon, Iris
 APPLICANT: Yaron, Avraham
 TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING ACTIVATION OF
 TITLE OF INVENTION: NF-kB
 FILE REFERENCE: 860098.427
 CURRENT APPLICATION NUMBER: US/09/832,161
 PRIOR FILING DATE: 2001-04-09
 PRIOR APPLICATION NUMBER: 09/210,060
 NUMBER OF SEQ ID NOS: 30
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 15
 LENGTH: 4230
 TYPE: DNA
 ORGANISM: Homo sapiens
 09-832-161-15
 Alignment Scores:
 Seq. No.: 6,89e-291 Length: 4230
 Score: 2384.50 Matches: 451
 Percent Similarity: 87.39% Conservative: 48
 Percent Local Similarity: 78.98% Mismatches: 41
 Identity Match: 78.59% Indels: 31
 Gaps: 7
 09-601-168B-2 (1-569) x US-09-832-161-15 (1-4230)
 1 MetAspProAlaGluValLeuGlnGluLysPheMetAsnSerSerGlu 20
 26 ATGGAGCCC---GACTCGGTGATTGAGGACCAAGACCATCGAGCTCATGTGTTCT----- 76
 21 ArgGluAspCysAsnAsndGlyGluProArgLysIleIleProGluLysAsnSerLeu 40
 77 -----GTGCCA----- 82
 41 ArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGlnThrValCysLeuAla 60
 83 AGGTCTTTGTGGTAGCTGCGCCAACTCGTGA-----GAGAGCATGTGCGCACTG 133
 61 Ser-----ThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGly 78
 134 AGTTGCTGTCAGAGCAAGTCCCGAGTGTCTGATGCTC-----CAGATAGTAATGGA 184
 79 ThrSerSerMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLys 98
 185 ACATCATCTGTGATCGTCTCCAGAAAGAGGCCCATCAGAGGAAACTATCAAAAGGAAAAA 244
 99 GluLeuCysValLysTyrPheGluGlnTrpSerGluSerAspGlnValGluPheValGlu 118
 245 GACTTGTGTAATAATTTTACAGATGGTCTGTAATCAGATCAAGTGGAAATTTGTGGAA 304
 119 HisLeuIleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysPro 138
 305 CATCTTAATTTCCAGAAATGTCTATATCAGCATGGACATATTAATCTTACCTGAAGCCC 364
 139 MetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGlu 158
 365 ATGTTCAGCGGAGCTTTATACCGCTTTACAGAGCAAGGCTTAGATCATCATAGCAGAA 424
 159 AsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGlu 178
 425 AACATTTCTTTGCTACCTGGATGCGAGGTCTCTGTGTGCGAGCAGAGCTGGTATGTAAGAA 494
 179 TrpTyrArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArg 198
 485 TGGCAGCGAGTATCTCAGAGGAATGCTTTGGAAGAAGCTGATTTGAACGATGTGTACGC 544
 199 ThrAspSerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLys 218
 545 ACTGATCCCTATGGAAGAGGACTTTCCAGAAAGAGAGGAGGTGGATCATGCTGTTTAAA 604

RESULT 5

219 AsnLysProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLys 238
 605 AACAGACCCACAGATGGC-----CTCCAAATTCATTTATAGTCAATATACCCAAAG 658
 239 IleIleGlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGln 258
 659 ATTATCCAGGATATAGAGACTATAGATCTTAACCTGGGGGTGTGGACGACACAACTTGCAG 718
 259 ArgIleHisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGln 278
 719 AGGATTCAGTGGCGCTCTGAAATAGTAAGGTGTCTACTGTTTACAGTACGATGATGAA 778
 279 LysIleValSerGlyLeuArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGlu 298
 779 AAAATTTATCAGTGGCTACGAGATAATCTTATTAGATATGGGATAAAACCAAGCTGGAA 838
 299 CysLysArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAspGluArg 318
 839 TGTTTGAAGAGTGTAAACAGGACACACAGGCTCTCTCTCTCTGTCGCAATGATGAGCGT 898
 319 ValIleIleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGlu 338
 899 GTCATTTGTAACCTGCTCTCAATTCACGGTGTAGAGTGTGGGATGTGAACACGGGTGAA 958
 339 MetLeuAsnThrLeuIleHisCysGluAlaValLeuHisLeuArgPheAsnAsnGly 358
 959 GTTCTTAAACATCTGATCCACCAATGAGGCTGTATTGCACCTACGCTTCAGCAATGGA 1018
 359 MetMetValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThr 378
 1019 CTGATGTCGACCTGTTTCCAAAGACCGCTCCATCTCTGTGTGGGACATGGCTTCGCGACC 1078
 379 AspIleThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValValAspPhe 398
 1079 GATCATCATTTACGGCGTGTCTGTTGCGCCACCGGCTGCGCTCAATGTAGTAGACTTT 1138
 399 AspAspLysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSer 418
 1139 GAGCACAAGTACATCGTGTCTCTCTCTGTCAGAGGACCATCAAAAGTCTGGAGCAGGC 1198
 419 ThrCysGlyPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyr 438
 1199 ACCTGTGAAATTTGTTCTGTTACTCTCAATGGCACAAGCGGGCATTTGCTGTCTCCAGTAC 1258
 439 ArgAspArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGlu 458
 1359 AGGATCGCGCTGTTGTTAGTGGATCATCAGATATATACCATTAGGCTCTGGGATTTGAA 1318
 459 CysGlyAlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPhe 478
 1319 TGTGGTGCCTGTTTAAAGATCTCTAGAGGACATGAAGAATTTGGTCCGATGTCATCCGCTTT 1378
 479 AspAsnLysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuVal 498
 1379 GATAACAAGAGGATTTGTCAGTGGGCTCTATGATGGGAAATTAAGTTTGGGACTTGCAA 1438
 499 AlaAlaLeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHis 518
 1439 GCTGCTCTTGACCTTCGACCCCGACAGACATTTGTGTTTGGCAGCATTTGGTGGAAAT 1498
 519 SerGlyArgValPheArgLeuGlnPheAspGluPheGlnIleValSerSerSerHisAsp 538
 1499 TCTGGACGTGTGTCTCGGCTCCAGTTTATGAGTTTTCAGATCATCAGCAGCTCCCATGAT 1558
 539 AspThrIleLeuIleTrpAspPheLeuAsnAspProAlaAlaGlnAlaGluProArg 558
 1559 GACATATTTTGAATTTGGGATTTCTTAAATGTGCTCCAGTGGCCAGATGAGACCGGT 1618
 559 SerProSerArgThrTyrThrTyrIleSerArg 569
 1619 TCTCCCTCCAGAACATACACTTACATCTCTAGA 1651

S-10-023-530-1

Sequence 1, Application US/10023530

Publication No. US20030007956A1

GENERAL INFORMATION:

APPLICANT: LEGRAIN, Pierre

APPLICANT: BENAROUS, Richard

APPLICANT: BLOT, Guillaume

APPLICANT: LASSOT, Irina

TITLE OF INVENTION: PROTEINS THAT INTERACT WITH BETA TrCP

FILE REFERENCE: B4717A

CURRENT APPLICATION NUMBER: US/10/023,530

CURRENT FILING DATE: 2002-04-22

PRIOR APPLICATION NUMBER: 60/256,276

PRIOR FILING DATE: 2000-12-18

NUMBER OF SEQ ID NOS: 30

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 657

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: Gene

LOCATION: (1)-(657)

OTHER INFORMATION: Beta TrCP

S-10-023-530-1

Alignment Scores:

Seq. No.:	2,71e-137	Length:	657
Score:	1163.00	Matches:	219
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	18.33%	Indels:	0
DB:	14	Gaps:	0

S-09-601-168B-2 (1-569) x US-10-023-530-1 (1-657)

NY	1	MetAspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGlu	20
NY	1	ATGGACCCGCGCGCGGTGTGCAAGAGAGAGGCACTCAAGTTTATGAATTCCTCAGAG	60
NY	21	ArgGluAspCysAsnAsnGlyGluProArgLysIleIleProGluLysAsnSerLeu	40
NY	61	AGAGAGAGACTGTAAATATGCGCAACCCCTAGAGAGATAATACCGAGAGAAATTCATT	120
NY	41	ArgGlnThrTyrrAsnSerCysAlaArgLeuCysLeuAsnGlnGluThrValCysLeuAla	60
NY	121	AGACAGACATACACAGACTGTCCGAGCTCTGCTTAAACCAAGAAACAGATGTTTASCA	180
NY	61	SerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSer	80
NY	181	AGCACTGCTATGAAGACTGAGAAATTGTGTGGCCAAACAAACATTCGCATGGCACITCC	240
NY	81	SerMetIleValProLysGlnArgLysLeuSerAlaSerTyrrGluLysGluLysLeu	100
NY	241	AGTATGATTGTGCCAAGCAACGAAACTCTCAGCAAGCTATGAAAGAGAAAGAACTG	300
NY	101	CysValLysTyrrPheGluGlnThrSerGluSerAspGlnValGluPheValGluHisLeu	120
NY	301	TGTGTCAAAATACCTTTGAGCAGTGGTTCAGAGTCAGATCAAGTGGAAATTTGTGAAACATCTT	360
NY	121	IleSerGlnMetCysHisTyrrGlnHisGlyHisIleAsnSerTyrrLeuLysProMetLeu	140
NY	361	ATATCCCAATGTGTCATTACCAACATGGGCACATPAACCTCGTATCTTAAACCTATGTTG	420
NY	141	GlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsnIle	160
NY	421	CAGAGAGATTTTCAATCTGCTTCGCAGCTCGGGATTCGATCATATCGCTGAGAACATT	480
NY	161	LeuSerTyrrLeuAspAlaLysSerLeuCysAlaAlaGluLeuValCysLysGluTrpTyrr	180
NY	481	CTGTCACTCTGGATGCCAAATCACTATGCTGTGCTGAACCTTGTGTGAAGAAATGGTAC	540
NY	181	ArgValThrSerAspGlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAsp	200

Db	541	CGAGTGACCTCTGATGGCATGCTGTGGAGAGGCTTATCGAGAGATGGTCAGACAGAT	600
Qy	201	SerLeuTrpArgGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrrLeuPheLysAsn	219
Db	601	TCTCTGTGGAGAGGCTGCGAGAACGAGAGAGATGGGCACAGTATTTATTCAAAC	657

RESULT 6

US-09-764-848-15
 ; Sequence 15, Application US/09764848
 ; Patent No. US20020077270A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT208
 ; CURRENT APPLICATION NUMBER: US/09/764,848
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 53
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 15
 ; LENGTH: 951
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (883)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (913)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (931)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (935)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (941)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; OTHER INFORMATION: n equals a,t,g, or c

US-09-764-848-15
 Alignment Scores:
 Seq. No.: 4,47e-136 Length: 951
 Score: 1155.50 Matches: 226
 Percent Similarity: 76.97% Conservative: 28
 Best Local Similarity: 68.48% Mismatches: 37
 Query Match: 38.09% Indels: 41
 DB: 9 Gaps: 3

Qy	15	PheMetAsnSerSerGluArgGluAspCysAsnAsnGlyGluProProArgLysIleIle	34
Db	59	TTCAGACACACTTCAGTTATGGAGATCAAAATGAAGATGAGTCCCAAG	109
Qy	35	ProGluLysAsnSerLeuArgGlnThrTyrrAsnSerCysAlaArgLeuCysLeuAsnGln	54
Db	110	-----AAAATACCTTTGGCAG-----	127
Qy	55	GluThrValCysLeuAlaSerThrAlaMetLysThrGluAsnCysValAlaLysThrLys	74
Db	127	-----	127
Qy	75	LeuAlaAsnGlyThrSerSerMetIleValProLysGlnArgLysLeuSerAlaSerTyrr	94
Db	128	ATAAGTAATGGAACATCATCTGTGATCGTCTCCAGAAAGAGGCCATCAGAGAACTAT	187
Qy	95	GluLysGluLysGluLeuCysValLysTyrrPheGluGlnTrpSerGluSerAspGlnVal	114
Db	188	CAAAAAGAAAAGACTTGTGTATTAATATTTTACCAGTGTCTGCAATCAGATCAAGTG	247
Qy	115	GluPheValGluHisLeuIleSerGlnMetCysHisTyrrGlnHisGlyHisIleAsnSer	134

b 248 GAATTGTGGACATCTTTATTTACGAATGTGTCAATATACGATGACATATTAATCTCT 307
Y 135 TyrLeuLysProMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAsp 154
b 308 TACCTGAAAGCCATGTTGACGGGACCTTTATTACCGCTTTACAGAGCAAGCGTTAGAT 367
Y 155 HisIleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeu 174
b 368 CACATAGCAGAAACATCTTTCGTACCTGGATGCCAGGTCTCTGTGTGAGCAGAGCTG 427
Y 175 ValCysLysGluTyrTyrArgValThrSerAspGlyMetLeuTyrLysLeuIleGlu 194
b 428 GTATGTAAAGAAATGACGAGGATGATCTCAGAAGGATGCTTTGGAAGAGCTGATTGAA 487
Y 195 ArgMetValArgThrAspSerLeuTyrArgGlyLeuAlaGluArgGlyTyrGlyGln 214
b 488 CGAATGGTACGCACTGATCCCTATGGAAGAGCACTTTCAGAAAGAGAGGTGGATCAG 547
Y 215 TyrLeuPheLysAsnLysProProAspGlyAsnAlaProProAsnSerPheTyrArgAla 234
b 548 TACCTGTTTAAAGAACAGACCCACAGATGCC-----CCTCCAAATTCATTTATAGTCA 601
Y 235 LeuTyrProLysIleIleGluAspIleGluThrIleGluSerAsnTyrArgCysGlyArg 254
b 602 TTATACCCAAAGATATCCAGATATAGACTATAGAACTTAAGTGGCGGTGGACGA 661
Y 255 HisSerLeuGlnArgIleHisCysArgSerGluThrSerLysGlyValTyrCysLeuGln 274
b 662 CACAACTTGCAGAGGATTCAGTGGCGCTCTGAAATAGTAAAGGTGCTACTGTTTACAG 721
Y 275 TyrAspAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTyrAspLys 294
b 722 TACGATGATGAAAAAATATCAGTGGCT-CGAGATAATCTATTAGATATGGATAAA 780
Y 295 AsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGln 314
b 781 ACCWGCCTGGAATGTTGAAAGTGTWACAGACACACAGCTCTGK-CTCTGTCTGACG 839
Y 315 TyrAspGluArgValIleIleThrGlySerSerAspSerThrValArgValTyrAspVal 334
b 840 TAIGATGAKCGGTGATTATKTWACTGGCTCTTCAGATTCACGGNAGAGAGTGGGATGTG 899
Y 335 AsnThrGlyGluMetLeuAsnThrLeuIle 344
b 900 ACACCGGTGAAAGNCTTAAACACATGTATC 929

RESULT 7

S-10-222-020-15

Sequence 15, Application US/10222020

Publication No. US2003017539A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PT208C2

CURRENT APPLICATION NUMBER: US/10/222,020

CURRENT FILING DATE: 2002-08-16

PRIORITY APPLICATION NUMBER: 10/116,016

PRIORITY FILING DATE: 2002-04-05

PRIORITY APPLICATION NUMBER: 09/764,848

PRIORITY FILING DATE: 2001-01-17

PRIORITY APPLICATION NUMBER: 60/179,065

PRIORITY FILING DATE: 2000-01-31

PRIORITY APPLICATION NUMBER: 60/180,628

PRIORITY FILING DATE: 2000-02-04

PRIORITY APPLICATION NUMBER: 60/214,886

PRIORITY FILING DATE: 2000-06-28

PRIORITY APPLICATION NUMBER: 60/217,487

PRIORITY FILING DATE: 2000-07-11

PRIORITY APPLICATION NUMBER: 60/225,758

PRIORITY FILING DATE: 2000-08-14

PRIORITY APPLICATION NUMBER: 60/220,963

PRIORITY FILING DATE: 2000-07-26

/ PRIOR APPLICATION NUMBER: 60/217,496
/ PRIOR FILING DATE: 2000-07-11
/ PRIOR APPLICATION NUMBER: 60/225,447
/ PRIOR FILING DATE: 2000-08-14
/ PRIOR APPLICATION NUMBER: 60/218,290
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: 60/225,757
/ PRIOR FILING DATE: 2000-08-14
/ PRIOR APPLICATION NUMBER: 60/226,868
/ PRIOR FILING DATE: 2000-08-22
/ PRIOR APPLICATION NUMBER: 60/216,647
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/225,267
/ PRIOR FILING DATE: 2000-08-14
/ PRIOR APPLICATION NUMBER: 60/216,880
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/225,270
/ PRIOR FILING DATE: 2000-08-14
/ PRIOR APPLICATION NUMBER: 60/251,869
/ PRIOR FILING DATE: 2000-12-08
/ PRIOR APPLICATION NUMBER: 60/235,834
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: 60/234,274
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: 60/234,223
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: 60/228,924
/ PRIOR FILING DATE: 2000-08-30
/ PRIOR APPLICATION NUMBER: 60/224,518
/ PRIOR FILING DATE: 2000-08-14
/ PRIOR APPLICATION NUMBER: 60/236,369
/ PRIOR FILING DATE: 2000-09-29
/ PRIOR APPLICATION NUMBER: 60/224,519
/ PRIOR FILING DATE: 2000-08-14
/ PRIOR APPLICATION NUMBER: 60/220,964
/ PRIOR FILING DATE: 2000-07-26
/ PRIOR APPLICATION NUMBER: 60/241,809
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/249,299
/ PRIOR FILING DATE: 2000-11-17
/ PRIOR APPLICATION NUMBER: 60/236,327
/ PRIOR FILING DATE: 2000-09-29
/ PRIOR APPLICATION NUMBER: 60/241,785
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/244,617
/ PRIOR FILING DATE: 2000-11-01
/ PRIOR APPLICATION NUMBER: 60/225,268
/ PRIOR FILING DATE: 2000-08-14
/ PRIOR APPLICATION NUMBER: 60/236,368
/ PRIOR FILING DATE: 2000-09-29
/ PRIOR APPLICATION NUMBER: 60/251,856
/ PRIOR FILING DATE: 2000-12-08
/ PRIOR APPLICATION NUMBER: 60/251,866
/ PRIOR FILING DATE: 2000-12-08
/ PRIOR APPLICATION NUMBER: 60/229,344
/ PRIOR FILING DATE: 2000-09-01
/ PRIOR APPLICATION NUMBER: 60/234,997
/ PRIOR FILING DATE: 2000-09-25
/ PRIOR APPLICATION NUMBER: 60/229,343
/ PRIOR FILING DATE: 2000-09-01
/ PRIOR APPLICATION NUMBER: 60/229,345
/ PRIOR FILING DATE: 2000-09-01
/ PRIOR APPLICATION NUMBER: 60/229,287
/ PRIOR FILING DATE: 2000-09-01
/ PRIOR APPLICATION NUMBER: 60/229,513
/ PRIOR FILING DATE: 2000-09-05
/ PRIOR APPLICATION NUMBER: 60/231,413
/ PRIOR FILING DATE: 2000-09-08
/ PRIOR APPLICATION NUMBER: 60/229,509
/ PRIOR FILING DATE: 2000-09-05
/ PRIOR APPLICATION NUMBER: 60/236,367
/ PRIOR FILING DATE: 2000-09-29
/ PRIOR APPLICATION NUMBER: 60/237,039

PRIOR FILING DATE: 2000-10-02
 PRIOR APPLICATION NUMBER: 60/237,038
 PRIOR FILING DATE: 2000-10-02
 PRIOR APPLICATION NUMBER: 60/236,370
 PRIOR FILING DATE: 2000-09-29
 PRIOR APPLICATION NUMBER: 60/236,802
 PRIOR FILING DATE: 2000-10-02
 PRIOR APPLICATION NUMBER: 60/237,037
 PRIOR FILING DATE: 2000-10-02
 PRIOR APPLICATION NUMBER: 60/237,040
 PRIOR FILING DATE: 2000-10-02
 PRIOR APPLICATION NUMBER: 60/240,960
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/239,935
 PRIOR FILING DATE: 2000-10-13
 PRIOR APPLICATION NUMBER: 60/239,937
 PRIOR FILING DATE: 2000-10-13
 PRIOR APPLICATION NUMBER: 60/241,787
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/246,474
 PRIOR FILING DATE: 2000-11-08
 PRIOR APPLICATION NUMBER: 60/246,532
 PRIOR FILING DATE: 2000-11-08
 PRIOR APPLICATION NUMBER: 60/249,216
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,210
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/226,681
 PRIOR FILING DATE: 2000-08-22
 PRIOR APPLICATION NUMBER: 60/225,759
 PRIOR FILING DATE: 2000-08-14
 PRIOR APPLICATION NUMBER: 60/225,213
 PRIOR FILING DATE: 2000-08-14
 PRIOR APPLICATION NUMBER: 60/227,182
 PRIOR FILING DATE: 2000-08-22
 PRIOR APPLICATION NUMBER: 60/225,214
 PRIOR FILING DATE: 2000-08-14
 PRIOR APPLICATION NUMBER: 60/235,836
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: 60/230,438
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/215,135
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: 60/225,266
 PRIOR FILING DATE: 2000-08-14
 PRIOR APPLICATION NUMBER: 60/249,218
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,208
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,213
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,212
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,207
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,245
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,244
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,217
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,211
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,215
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,264
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,214
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/249,297
 PRIOR FILING DATE: 2000-11-17
 PRIOR APPLICATION NUMBER: 60/232,400
 PRIOR FILING DATE: 2000-09-14

; PRIOR APPLICATION NUMBER: 60/231,242
 ; PRIOR FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: 60/232,081
 ; PRIOR FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: 60/232,080
 ; PRIOR FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: 60/231,414
 ; PRIOR FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: 60/231,244
 ; PRIOR FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: 60/233,064
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: 60/233,063
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: 60/232,397
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: 60/232,399
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: 60/232,401
 ; PRIOR FILING DATE: 2000-09-14
 ; PRIOR APPLICATION NUMBER: 60/241,808
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/241,826
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/241,786
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/241,221
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/246,475
 ; PRIOR FILING DATE: 2000-11-08

Alignment Scores:

Pred. No.:	4,47e-136	Length:	951
Score:	1155.50	Matches:	226
Percent Similarity:	76.97%	Conservative:	28
Best Local Similarity:	68.48%	Mismatches:	37
Query Match:	38.09%	Indels:	41
DB:	12	Gaps:	3

US-09-601-168B-2 (1-569) X US-10-222-020-15 (1-951)

QY	15	PheMetAsnSerSerGluArgGluAspCysAsnAsnGlyGluProProArgLysIle	34
Db	59	TTCCGAGAACTTCAGTTATGGAAGATCAAATGAAGATGAGTCCCAAG-	109
QY	35	ProGluLysAsnSerLeuArgGluThrTyrrAsnSerCysAlaArgLeuCysLeuAsnGln	54
Db	110	-----AAAATACCTTTGGCAG-	127
QY	55	GluThrValCysLeuAlaSerThrAlaMetLysThrGluAsnCysValAlaLysThrLys	74
Db	127	-----	127
QY	75	LeuAlaAsnGlyThrSerSerMetIleValProLysGlnArgLysLeuSerAlaSerTyr	94
Db	128	ATAAGTAATGGAACATCATCTGTGATTAATAATTTTGCACCATGGTCTGAATCAGATCAAGTG	187
QY	95	GluLysGluLysGluLeuCysValLysTyrPheGluGlnTrpSerGluSerAspGlnVal	114
Db	188	CAAAAAGAAAAGACTTGTGTATTAAATATTTTGCACCATGGTCTGAATCAGATCAAGTG	247
QY	115	GluPheValGluHisLeuIleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSer	134
Db	248	GAATTTGTGGAACTCTTATTTCACGAATGTGTCATTATCAGCATGGACATATTAACTCT	307
QY	135	TyrLeuLysProMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAsp	154
Db	308	TACCTGAAGCCCATGTTCCAGCGGACTTTATTACCGCTTTACAGAGCAGGCTTAGAT	367
QY	155	HisIleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeu	174
Db	368	CACATAGCAGAAAACATCTTTCTGCTGATCCAGGTCTCTGTGTGACAGCAGAGCTG	427

175 ValCysLysGluTyrArgValThrSerAspGlyMetLeuTyrPylsLysLeuIleGlu 194
 b 428 GTATGTAAGATGCGAGGAGTGTCTCAGAGGAATGCTTTGSAAGAGCTGATTGAA 487
 y 195 ArgMetValArgThrAspSerLeuTyrArgGlyLeuAlaGluArgGlyTyrGlyGln 214
 b 488 CGAATGGTACGCACTGATCCCTATGGAAGAGCTTTTCAGAAAGAGAGGTGGATCAG 547
 y 215 TyrLeuPheLysAsnLysProPheAspGlyAsnAlaProPheAsnSerPheTyrArgAla 234
 b 548 TACCTGTTTAAAGACAGCCACAGATGCT-----CCTCAAAATTCATTTATAGTCA 601
 y 235 LeuTyrProLysIleGlnAspIleGluThrIleGluSerAsnTyrArgCysGlyArg 254
 b 602 TTATACCCAAAGATTATCCAGGATATAGAGTCTGAAATAGTAAAGGTGCTACTGTTTACAG 661
 y 255 HisSerLeuGlnArgIleHisCysArgSerGluThrSerLysGlyValTyrCysLeuGln 274
 b 662 CAAACCTTCAGAGATTTCAGTCCGCTCTGAAATAGTAAAGGTGCTACTGTTTACAG 721
 y 275 TyrAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTyrAspLys 294
 b 722 TAGCATGATGAAAAAATTTATCAGTGGCT-CGAGATATTTCTAAGATATAGGATGAA 780
 y 295 AsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGln 314
 b 840 TATGATGAKCCGTGATTCTTACTGCTTACAGGACACACAGCTCTGK-CTCTGCTGCAG 839
 y 315 TyrAspGluArgValIleThrGlySerSerAspSerThrValArgValTyrAspVal 334
 b 840 TATGATGAKCCGTGATTCTTACTGCTTACAGGACACACAGCTCTGK-CTCTGCTGCAG 899
 y 335 AsnThrGlyGluMetLeuAsnThrLeuIle 344
 b 900 ACACCGTGAAGAGTCTTAACACATTGATC 929
 RESULT 8
 S-10-116-016-15
 Sequence 15, Application US/10116016
 Publication No. US20030054379A1
 GENERAL INFORMATION:
 APPLICANT: Rosen et al.
 TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 FILE REFERENCE: PT208C1
 CURRENT APPLICATION NUMBER: US/10/116,016
 PRIORITY FILING DATE: 2002-04-05
 Prior Application removed - See File Wrapper or Palm
 NUMBER OF SEQ ID NOS: 53
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 15
 LENGTH: 951
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (883)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: misc feature
 LOCATION: (913)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: misc feature
 LOCATION: (931)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: misc feature
 LOCATION: (935)
 OTHER INFORMATION: n equals a,t,g, or c
 NAME/KEY: misc feature
 LOCATION: (941)
 OTHER INFORMATION: n equals a,t,g, or c
 S-10-116-016-15
 Alignment Scores:
 red. No.: 4.47e-136 Length: 951

Score: 115.50 Matches: 226
 Percent Similarity: 76.97% Conservative: 28
 Best Local Similarity: 68.48% Mismatches: 37
 Query Match: 38.09% Indels: 41
 DB: 14 Gaps: 3
 US-09-601-168B-2 (1-569) x US-10-116-016-15 (1-951)
 QY 15 PheMetAsnSerSerGluArgGluAspCysAsnAsnGlyGluProArgLysIleIle 34
 Db 59 TTCAGAACACTTCAGTTATGGAAGATCAAAATGAAGATGAGTCCCAAG----- 109
 QY 35 ProGluLysAsnSerLeuArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGln 54
 Db 110 -----AAAATACTCTTTGGCAG----- 127
 QY 55 GluThrValCysLeuAlaSerThrAlaMetLysThrGluAsnCysValAlaLysThrLys 74
 Db 127 ----- 127
 QY 75 LeuAlaAsnGlyThrSerSerMetIleValProLysGlnArgLysLeuSerAlaSerTyr 94
 Db 128 ATAAAGTAATGGAACATCATCTGTGATCGTCTCCAGAAAGAGGCCATCAGAGGAACTAT 187
 QY 95 GluLysGluLysGluLeuCysValLysTyrPheGluGlnTyrSerGluSerAspGlnVal 114
 Db 188 CAAAAGAAAAAGAACTTGTGTATTAATATTTTGCACAGTGGTCTGAATCAGATCAAGTG 247
 QY 115 GluPheValGluHisLeuIleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSer 134
 Db 248 GAATTTGTGGACATCTTATTTACGNAATGTGTCAATATCAGCATGCGATATTAACTCT 307
 QY 135 TyrLeuLysProMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgLysLeuAsp 154
 Db 308 TACCTGAAGCCCATGTTGCGAGCGGACTTTTATACCGCTTTACGAGCAAGCTTAGAT 367
 QY 155 HisIleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeu 174
 Db 368 CACATAGCAGAAACATCTTTTGTACCTGGATGCCAGGTCTCTGTGCGCAGCAGCTG 427
 QY 175 ValCysLysGluTyrArgValThrSerAspGlyMetLeuTyrPylsLysLeuIleGlu 194
 Db 428 GTATGTAAGAAATGCGCAGGAGTGTCTCAGAAAGAAATGCTTTGGAAGAAAGCTGATTGAA 487
 QY 195 ArgMetValArgThrAspSerLeuTyrArgGlyLeuAlaGluArgArgGlyTyrGlyGln 214
 Db 488 CGAATGGTACGCACTGATCCCTATGGAAGAGCTTTTCAGAAAGAGAGGTGGATCAG 547
 QY 215 TyrLeuPheLysAsnLysProPheAspGlyAsnAlaProPheAsnSerPheTyrArgAla 234
 Db 548 TACCTGTTTAAAGACAGCCACAGATGCT-----CCTCAAAATTCATTTATAGTCA 601
 QY 235 LeuTyrProLysIleGlnAspIleGluThrIleGluSerAsnTyrArgCysGlyArg 254
 Db 602 TTATACCCAAAGATTATCCAGGATATAGAGTCTGAAATCAGTGGCGGTGGAGCA 661
 QY 255 HisSerLeuGlnArgIleHisCysArgSerGluThrSerLysGlyValTyrCysLeuGln 274
 Db 662 CAAACCTTCAGAGATTTCAGTCCGCTCTGAAATAGTAAAGGTGCTACTGTTTACAG 721
 QY 275 TyrAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTyrAspLys 294
 Db 722 TAGCATGATGAAAAAATTTATCAGTGGCT-CGAGATATTTCTAAGATATAGGATGAA 780
 QY 295 AsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGln 314
 Db 781 ACCGCGCTGGATGTTTGAAGGTGTTWACAGGACACACAGCTCTGK-CTCTGCTGCAG 839
 QY 315 TyrAspGluArgValIleThrGlySerSerAspSerThrValArgValTyrAspVal 334
 Db 840 TATGATGAKCCGTGATTCTTACTGCTTACAGGACACACAGCTCTGK-CTCTGCTGCAG 899
 QY 335 AsnThrGlyGluMetLeuAsnThrLeuIle 344

b 900 ACACCGGTGAAGACGCTTAACACATTGATC 929

RESULT 9

S-09-864-761-2778/c

Sequence 2778, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Shazron G.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aeomica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

PRIOR FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/774,203

PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

SEQ ID NO 2778

LENGTH: 479

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AP000252.1

OTHER INFORMATION: EXPRESSED IN HELI00, SIGNAL = 3.4

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.6

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.1

OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.8

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.75

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.3

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7

Alignment Scores:

Pred. No.: 2,13e-77 Length: 479
Score: 689.00 Matches: 140
Percent Similarity: 90.12% Conservative: 6
Best Local Similarity: 86.42% Mismatches: 13
Query Match: 22.71% Indels: 3
DB: 9 Gaps: 1

US-09-601-168B-2 (1-569) x US-09-864-761-2778 (1-479)

QY 360 MetValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerP-oThrAsp 379
DB 478 ATGGTGAACCTGTTCCAGGACCGCTCCATTGCTGTGTGGGACATGCTCTTCGACCGAT 419
QY 380 IleThrLeuArgValLeuValGlyHisArgAlaValAsnValValAspPheAsp 399
DB 418 ATCACTTTTACGTGGTGTCTGCTGCTCCACTGAGCTGCTCAACCTAGTAGACTTTGAC 359
QY 400 AspLysTyrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSerThr 419
DB 358 GAGAAATACATCGTGTCTGCTCTGGTGACAGGACCATCAAGCTCTGGAGCATGACGCC 299
QY 420 CysGluPheValArgThrLeuAsnGlyHisLysArgGlyIleAla-CysLeuGlnTyrAr 439
DB 298 TGTGAATTGCTCGTACTCTCAATGGGCAACGCA-----GTCTTGTCTCCAGTACAG 245
QY 439 GAspArgLeuValSerGlySerSerAspAsnThrIleArgLeuTrpAspIleGluCy 459
DB 244 GGATGGGCTGGTGTAGTGGATCATCAGATCATACCATAGGCTCTTGATATTGATG 185
QY 459 sGlyAlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIleArgPheAs 479
DB 184 TGGCTCTCTGTTTAAAGAGTCTTAGAGGACATGAGAAATTGTCGCGATGCATCCAGTTGA 125
QY 479 pAsnLysArgIleValSerGlyAlaTyrAspGlyLysIleLysValTrpAspLeuValAl 499
DB 124 TAAACAAGAGGATGTCAGCCGGGCTATGATGGGAATAATTAAAGTTAGGACTTGCACG 65
QY 499 aAlaLeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSe 519
DB 64 CGCTCTTGACCTCGAGCCGCCAGCAGCAGCATTTGTTCGCACATTGTTCGCACTTGTGGA 5
QY 519 rGly 520
DB 4 TGA 1

RESULT 10

US-09-213-888-20

Sequence 20, Application US/09213888A

Patent No. US20020164683A1

GENERAL INFORMATION:

APPLICANT: Gurney, Mark E.

APPLICANT: Li, Jinhe

APPLICANT: Pauley, Adele M.

APPLICANT: Pharmacia & Upjohn Company

TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that

TITLE OF INVENTION: Encode Them

FILE REFERENCE: 6142

CURRENT APPLICATION NUMBER: US/09/213,888A

CURRENT FILING DATE: 1998-12-17

NUMBER OF SEQ ID NOS: 27

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 20

LENGTH: 1881

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: 6 myc tagged

OTHER INFORMATION: homo sapiens

US-09-213-888-20

Alignment Scores:

Pred. No.: 9.05e-71 Length: 1881


```

Db      1031  GGTGGAGTATGGTTCATCACAAATGAGGGCAACATCATCATCTAGTGGATCTACAGATCGG 1080
Qy      328   ThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuLeuHisHisCys 347
Db      1081  ACATCTCAAGTGTGGAAATGCAGAGACTGGAGAAATGTATACACACTTATATGGGCATACT 1140
Qy      348   GluAlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSerLysAspArg 367
Db      1141  TCCACTGTGGCGTGTATGCATCTTCATGAAAAAGAGTTGTAGCGGTCTCGAGATGCC 1200
Qy      368   SerIleAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgArgValLeuVal 387
Db      1201  ACTCTTAGGGTTGGGAATTTGAGACGCCAGCTGTTTA-----CATGTTTTCATG 1251
Qy      388   GlyHisArgAlaAlaValAsnValAspPheAspAspLysTyrlIleValSerAlaSer 407
Db      1252  GGTCAATGTCGAGCAGTCCGCTGTGTTCAATATGATGCGAGGAGGGTGTGTAGTGAGCA 1311
Qy      408   GlyAspArgThrIleLysValTrpAsnThrSerThrCysGluPheValArgThrLeuAsn 427
Db      1312  TATGATTTATGGTAAAGGTGGGATCCAGAGACTGAAACCTGTCTACACACGTTCCAG 1371
Qy      428   GlyHisLysArgGlyIleAlaCysLeuGlnTyraArgAspArgLeuValValSerGlySer 447
Db      1372  GGGCATACTAATAGAGTCTATTCAATTACAGTTTGATGGTATCCATGTTGGTGGATCT 1431
Qy      448   SerAspAsnThrIleArgLeuTrpAspIleGluCysGlyAlaCysLeuArgValLeuGlu 467
Db      1432  CTTGATACATCCATCCGCTGTTGGGATGTGGAGACAGGAATTCATTCATCACACGTTAAC 1491
Qy      468   GlyHisGluGluLeuValArgCysIleArgPheAspAsnLysArgIleValSerGlyAla 487
Db      1492  GGGCACACAGTCGTTTAAACAAGTGAATGGAATCAAAAGACAATATTCCTCTGGGAAT 1551
Qy      488   TyrAspGlyLysIleLysValTrpAspLeuValAlaAlaLeuAspProArgAlaProAla 507
Db      1552  GCAGATCTCAGATTAAATCTCGGATATCAAAACAGGA----- 1590
Qy      508   GlyThrLeuCysLeuArgThrLeuVal-----GluHisSerGlyArgValPheArg 524
Db      1591  -----CAGTGTTTTACAAACATTCGCAAGGTCCCAACAGCATCAGAGTGTGTGACCTGT 1644
Qy      525   LeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThrIleLeuIleTrp 544
Db      1645  TTACAGTTTCAACAAGAAGACTTTGTAATTACCACTCAGATGATGGAAGCTGTAATACTATGG 1704
Qy      545   Asp 545
Db      1705   GAC 1707

RESULT 11
US-09-328-877A-20
; Sequence 20, Application US/09328877A
; Patent No. US20020177187A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Li, Jinhe
; APPLICANT: Pauley, Adele M.
; APPLICANT: Pharmacia & Upjohn Company
; TITLE OF INVENTION: Human Sel-10 Polypeptides and Polynucleotides that
; TITLE OF INVENTION: Encode Them
; FILE REFERENCE: 6142
; CURRENT APPLICATION NUMBER: US/09/328,877A
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 1881
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 6 myc tagged

```

OTHER INFORMATION: homo sapiens
US-09-328-877A-20

Alignment Scores:			
Pred. No.:	9,05e-71	Length:	1881
Score:	644.50	Matches:	169
Percent Similarity:	48.48%	Conservative:	103
Best Local Similarity:	30.12%	Mismatches:	222
Query Match:	21.24%	Indels:	67
DB:	10	Gaps:	15
US-09-601-168B-2 (1-569) x US-09-328-877A-20 (1-1881):			
QY	8	LeuGlnGluLysAlaLeuLysPheMetAsnSerSerGluArgGluAspCysAsnAsnGly	27
DB	157	ATGGAGCAAAAGCTCATTTCTGAAGAGGACTTGAATGAAATGGAGAGCTTGGGGACCTC	216
QY	28	GluProProArgLysIleProGluLys-----AsnSerLeuArgGln-----Thr	43
DB	217	ACCATGGAGCAAAAGCTCATTTCTGAAGAGGACTTGAATTCATGAAAGAAAGAGTTGGAC	276
QY	44	TyrAsnSerCysAlaArgLeuCysLeuAsnGlnGlnThrValCysLeuAlaSerThrAla	63
DB	277	CAATGTTCTGAGCTCGCTCTTTCTTGGAAAGAAACCATGCAAGTCTCAGAAATAT	336
QY	64	MetLysThrGluAsnCysVal-----AlaLysThrLys	74
DB	337	ACAAGTACCACCTGGGCTTGTACCATGTTTCAGCAACACCAACAACTTTTGGGGACCTCAGA	396
QY	75	LeuAlaAsnGlyThrSerSerMetIleValProLysGlnArgLysLeuSerAlaSerTyr	94
DB	397	GCAGCCAAATGGCCCAAGG-----CAACAACACGACCGCAATATACATCTGTC	441
QY	95	GluLysGluLysGluLeu-----CysValLysTyrPheGluGlnThrSerGluSerAsp	112
DB	442	CAGCCACCTACAGGCTCCAGGAATGGCTAAATGTTTCAGAGCTGGAGTGGACAGAG	501
QY	113	GlnValGluPheValGluHisLeuSerGlnMetCysHisTyrGlnHisGlyHisIle	132
DB	502	AAATGCTTGTCTTAGATGAACATTCATATGATGTTGTGAACCAACACCAAGTAAACATATG	561
QY	133	AsnSerTyrLeuLysProMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGly	152
DB	562	ATGCAAGTATAGAACCCAGTTTCAACAGACTTCATTTCTGCTCCCT-----	612
QY	153	LeuAspHisIleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAlaAla	172
DB	613	---AAAGAGTTGGCACTCTATGTGCTTTTCATTCCTGGAAACCAAGACCTGCTACAAGCA	669
QY	173	GluLeuValCysLysGluTyrArgValThrSerAspGlyMetLeuTyrLysLeu	192
DB	670	GTCACACATGCTGCTACTGAGAAATTTGGCTGAAGCAACCTTCTCTGGAGAGAGAAA	729
QY	193	IleGluArgMetValArgThrAspSerLeuTyrPheGlyLeuAlaGlu-----	208
DB	730	TGCAAGAA-----GAGGGGATTGATGAACCATTCGCATC	765
QY	209	---ArgArgGlyTyrGlyGlnTyrLeuPheLysAsnLysProProAspGlyAsnAlaPro	227
DB	766	AAGAGAAGAAA-----GTAATAAACCCAGGTTTCATACACAGTCCA	807
QY	228	ProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleGluThrIleGlu	247
DB	808	TGAAAGAGTGCATAC-----ATCAGACAG-----CACAGAATTGAT	843
QY	248	SerAsnTyrArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSerGluThrSer	267
DB	844	ACTAACTGAGCGGAGGAGAACTCAATCTCCTAAGGTG---CTGAAAGACATGATGAT	900
QY	268	LysGlyValTyrCysLeuGlnTyrAspGlnLysIleValSerGlyLeuArgAspAsn	287
DB	901	CATGTGATCACAATGTTTGTGGTAAACCAATAGTATTGTTCTGTATGACAAC	960

QY	288	ThrIleLysIleTyrPheAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThr	307
DB	961	ACTTTAAAGATTTGGTTCAGCAGTCCAGGCAAAATGTCTGAGAACATTTAGTGGACATACA	1020
QY	308	GlySerValLeuCysLeuGlnTyrAspGluArgValIleIleThrGlySerSerAspSer	327
DB	1021	GGTGGAGTATGGTTCATCACAAATCAGGACAAATCATCATTTAGTGGATCTACAGATCGG	1080
QY	328	ThrValArgValTyrPheValAsnThrGlyGluMetLeuAsnThrLeuIleHisCys	347
DB	1081	ACACTCAAGTGTGGAATTCAGAGACTGGAGAAATGTATACACACCTTATATGGCACTACT	1140
QY	348	GluAlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSerLysAspArg	367
DB	1141	TCCACTGTGCTGTATGTCATCTTCATGAAAAAGAGTTGTTAGCGGTCTTCGAGATGCC	1200
QY	368	SerIleAlaValTyrPheMetAlaSerProThrAspIleThrLeuArgValLeuVal	387
DB	1201	ACTCTTAGGGTTCGGGATATTTGACACAGCCAGTGTGTTA-----CATGTTTGTATG	1251
QY	388	GlyHisArgAlaValAsnValValAspPheAspAspLysTyrIleValSerAlaSer	407
DB	1252	GGTCATGTTGGAGCAGTCCGCTGTGTTCAATATGATGGCAGGAGGTTGTTAGTGAGCA	1311
QY	408	GlyAspArgThrIleLysValTyrAsnThrSerThrCysGluPheValArgThrLeuAsn	427
DB	1312	TATGATTTTATGTTAAAGTGTGGATCCAGAGACTGAAACCTGTCTACACACGTTGCAG	1371
QY	428	GlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAspArgLeuValSerGlySer	447
DB	1372	GGGCATACCTAATAGAGTCTATTTCATACAGTTTGTATGATGATCATCATGTTGATGAT	1431
QY	448	SerAspAsnThrIleArgLeuTyrAspIleGluCysGlyAlaCysLeuArgValLeuGlu	467
DB	1432	CTTGATACATCCATCCGTTGTTGGATGTGGAGACAGGGAATTCATTCACACGTTAAACA	1491
QY	468	GlyHisGluGluLeuValArgCysIleArgPheAspAsnLysArgIleValSerGlyAla	487
DB	1492	GGGACACCACTGTTTAAACAGTGAATGGAATGGAATCAAAACAAATATTTCTTCTCGGAAT	1551
QY	488	TyrAspGlyLysIleLysValTyrAspLeuValAlaAlaLeuAspProAlaProAla	507
DB	1552	GCAGATTCACAGTTAAATCTGGGATATCAAAACAGGA-----	1590
QY	508	GlyThrLeuCysLeuArgThrLeuVal-----GluHisSerGlyArgValPheArg	524
DB	1591	-----CAGTGTTCACAAACATTTGCAAGGTCCCAAGCATCAGAGTGTGTGACCTGT	1644
QY	525	LeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThrIleLeuIleTyr	544
DB	1645	TTCAGTTCACAGAACTTTGTAATTCACGCTCAGATGATGGAACTGTAAACTATGG	1704
QY	545	Asp 545	
DB	1705	GAC 1707	

RESULT 12
US-10-245-618-13
; Sequence 13, Application US/10245618
; Publication No. US20030143582A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven
; APPLICANT: Strohmaier, Helmo
; APPLICANT: Spruck, Charles
; APPLICANT: Sangfelt, Olle
; TITLE OF INVENTION: HCDCA MODULATES CYCLIN E DEGRADATION
; FILE REFERENCE: TSRI 779.2
; CURRENT APPLICATION NUMBER: US/10/245,618
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/404,116
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/322,947
; PRIOR FILING DATE: 2001-09-14

NUMBER OF SEQ ID NOS: 48

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 13

LENGTH: 1620

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthesized

US-10-245-618-13

Alignment Scores:

red. No.: 2.63e-70 Length: 1620
Score: 640.00 Matches: 167
Percent Similarity: 47.79% Conservative: 92
Best Local Similarity: 30.81% Mismatches: 211
Query Match: 21.09% Indels: 72
Gaps: 15

S-09-601-168b-2 (1-569) x US-10-245-618-13 (1-1620)

14 LysPheMetAsnSerSerGluArgGluAspCysAsnAsnGlyGluProProArgLysIle 33
10 AAGTTGGACCATGGTCTCGAGTCCGCTCTTTCTTTGGGAAGAAACCATGCAGATC 69
34 IleProGluLysAsnSerLeuArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsn 53
70 -----TCGAATATACAAGTACCCTGGGCTT----- 96
54 GlnGluThrValCysLeuAlaSerThrAlaMetLysThrGluAsnCysValAlaLysThr 73
97 -----GTACCATGTTACGCA-----ACACCAACAACCTTTGGGACCTC 135
74 LysLeuAlaAsnGlyThrSerSerMetIleValProLysGlnArgLysLeuSerAlaSer 93
136 AGACGAGCCAAATGGCCAAAGG-----CAACAACGACCCGAATTTACATCT 180
94 TyrGluLysGluLysGluLeu-----CysValLysTyrPheGlnGlnThrPheSerGluSer 111
181 GTCCAGCCACCTACAGGCTCCAGGAATGGCTAAATGTTTCAGAGCTGAGTGGACCA 240
112 AspGlnValGluPheValGluHisLeuSerGlnMetCysHisTyrGlnHisGlyHis 131
241 GAGAAATTCGCTGCTTTAGATGAACCTCATTCATAGTTGTGAACCAACACCAAGTAAACAT 300
132 IleAsnSerTyrLeuLysProMetLeuGlnArgAspPheIleThrAlaLeuProAlaArg 151
301 ATGATGCAAGTCATAGACCCAGTTTCAACGAGACTTCATTTCATTCCTCCCT----- 354
152 GlyLeuAspHisIleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAla 171
355 -----AAAGAGTTGGCACTCTATGTGCTTTTCATTCCTGGAACCCAAAGACCTGCTACAA 408
172 AlaGluLeuValCysLysGluTyrPheArgValThrSerAspGlyMetLeuTyrLysLys 191
409 GCAGCTCAGACATGCTGCTACTGGAATTTGGCTGAAGCAACCTTCTCTGAGAGAG 468
192 LeuIleGluArgMetValArgThrAspSerLeuPheArgGlyLeuAlaGlu----- 208
469 AAATGCAAGAA-----GAGGGGATTTGATGAACCATTCGCAC 504
209 -----ArgArgGlyTyrGlyGlnTyrLeuPheLysAsnLysProProAspGlyAsnAla 226
505 ATCAGAGAGAA-----GTAATAAACCCAGGTTTCATACACAGT 546
227 ProProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleGluThrIle 246
547 CCATGGAAGTGCATAC-----ATCAGACAG-----CACAGAAATT 582
247 GluSerAsnTyrArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSerGluThr 266
583 GATACTAATGAGCGGAGGAGCACTCAATCTCCTAAGGTG---CTGAAAGACATGAT 639
267 SerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLeuArgAsp 286

Db 640 GATCATGTGATCATGCTTACAGTTTCTGTGTGTAACCGAATAGTTAGTGGTCTCTATGAC 699
Qy 287 AsnThrIleLysIleTyrAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHis 306
Db 700 AACACATTAAAGTTTGGTCAGCAGTCAAGGCAAAATGTCTGAGAACATTAGTGGACAT 759
Qy 307 ThrGlySerValLeuCysLeuGlnTyrAspGluArgValIleIleThrGlySerSerAsp 326
Db 760 ACAGTGGAGTATGGTCATCACAATGAGAGACACATCATCATTAGTGGATCTACAGAT 819
Qy 327 SerThrValArgValTyrAspValAsnThrGlyGluMetLeuAsnThrLeuIleHis 346
Db 820 CGGACACTCAAGTGTGGAATCGAGAGACTGGAGAATGTATACACACCTTATATGGCAT 879
Qy 347 CysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSerLysAsp 366
Db 880 ACTTCCACTGTGCGTGTATGCATCTTCATGAAAAAGAGTGTGTAGCGGTTCTCGAGAT 939
Qy 367 ArgSerIleAlaValTyrAspMetAlaSerProThrAspIleThrLeuArgArgValLeu 386
Db 940 GCCACTCTTAGGGTTGGGATATTGAGACAGAGCCAGTGTTA-----CATGTTTG 990
Qy 387 ValGlyHisArgAlaValAsnValValAspPheAspAspLysTyrIleValSerAla 406
Db 991 ATGGGTCACTGTTCCAGCAGTCCGCTGTGTTCAATATGATGGCAGGAGGTTGTAGTGA 1050
Qy 407 SerGlyAspArgThrIleLysValTyrAsnThrSerThrCysGluPheValArgThrLeu 426
Db 1051 GCATATGATTTTATGATAAGGTGTGGATCCAGAGACTGAAACCTGCTACACAGTTG 1110
Qy 427 AsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAspArgLeuValSerGly 446
Db 1111 CAGGGCCACTACTAATAGACTCTATTCAATACAGTTTGTATGGTATCCATGTGGTGTAGTGA 1170
Qy 447 SerSerAspAsnThrIleArgLeuTyrAspIleGluCysGlyValCysLeuArgValLeu 466
Db 1171 TCCTTGATACATCAATCCGCTGTTGGGATGGGAGACAGGGAATTCATTACACAGTTA 1230
Qy 467 GluGlyHisGluLeuValArgCysIleArgPheAspAsnLysArgIleValSerGly 486
Db 1231 ACAGGGCAGCAGTCGTAAACAGTGAATGGAACCTCAAGACAAATATCTTGTCTCTGGG 1290
Qy 487 AlaTyrAspGlyLysIleLysValTyrAspLeuValAlaAlaLeuAspProArgAlaPro 506
Db 1291 AATGCAGATTCTACAGTTAAATCTGGGATATCAAAACAGGA----- 1332
Qy 507 AlaGlyThrLeuCysLeuArgThrLeuVal-----GluHisSerGlyArgValPhe 523
Db 1333 -----CAGTGTTTACAAACATTGCAAGGTCCCAACAGCATCAGAGTGTCTGACC 1383
Qy 524 ArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThrIleLeuIle 543
Db 1384 TGTTACAGTTCAACAGAACTTTGTANTTACCACTCAGATGATGGAACTGTAAAACTA 1443
Qy 544 TrpAsp 545
Db 1444 TGGGAC 1449

RESULT 13

US-10-245-618-7

; Sequence 7, Application US/10245618

; Publication No. US20030143582A1

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven

; APPLICANT: Strohmaier, Heimo

; APPLICANT: Spruck, Charles

; APPLICANT: Sangfelt, Olle

; TITLE OF INVENTION: HCD04 MODULATES CYCLIN E DEGRADATION

; FILE REFERENCE: TSRI 779.2

; CURRENT APPLICATION NUMBER: US/10/245,618

; CURRENT FILING DATE: 2002-09-16

; PRIOR APPLICATION NUMBER: US 60/404,116

622	Db	GATCTAACTCTGGAGCGAGAGAACTCAAATCTCTTAGGGTG---CTGAAAGGACATGAT	678
267	QY	SerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLeuArgAsp	286
679	Db	GATCATGTGATCATGCTTACAGTTTGTGGTAACCGAATAGTAGTGGTCTTGATGAC	738
287	QY	AsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHis	306
739	Db	AACACTTTAAAGATTGGTCAGCAGTCACAGCAAAATCTTGAGAACATTAAGTGGGACAT	798
307	QY	ThrGlySerValLeuCysLeuGlnTyrAspGluArgValIleIleThrGlySerSerAsp	326
799	Db	ACAGGTGGATGTGGTCATCACAAATGAGACACACATCATTTAGTGGATCTACAGAT	858
327	QY	SerThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisHis	346
859	Db	CGGACACTCAAAGTGTGGAATGCAGAGATCGGAGAAATGTATACACACCTTATATGGGCAT	918
347	QY	CysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSerLysAsp	366
919	Db	ACTTCCACTGTGGGTGTATGCACTTTCATGAAAAAGAGTTGTAGCGGTTCTCGAGAT	978
367	QY	ArgSerIleAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgArgValLeu	386
979	Db	GCCCACTCTTAGGGTTTGGGATATGAGACAGCCAGTGTTA-----CATGTTTGT	1029
387	QY	ValGlyHisArgAlaAlaValAsnValValAspPheAsnAspLysTyrIleValSerAla	406
1030	Db	ATGGGTATGTTCAGCAGTCCCGCTGTGTTCAATATGATGGCAGAGGGTTGTAGTGGGA	1089
407	QY	SerGlyAspArgThrIleLysValTrpAsnThrSerThrCysGluPheValArgThrLeu	426
1090	Db	GCATATGATTTATGCTTAAAGGTGGGATCCAGAGACTGAAACCTGCTACACACGTTG	1149
427	QY	AsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAspArgLeuValValSerGly	446
1150	Db	CAGGGGCATACTAATAGAGTCTATTCATTACAGTTTGATGTATCCATGTGGTGAGTGA	1209
447	QY	SerSerAspAsnThrIleArgLeuTrpAspIleGluCysGlyAlaCysLeuArgValLeu	466
1210	Db	TCTCTTGATCATCAATCCGTGTGGATGTGGACAGACGGAATTCATTACACAGTTA	1269
467	QY	GluGlyHisGluGluLeuValArgCysIleArgPheAspAsnLysArgIleValSerGly	486
1270	Db	ACAGGGACCCAGTCGTAAACAAGTGAATGGAACCAAGACAAATATTCTTGTCTCTGGG	1329
487	QY	AlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAlaLeuAspProArgAlaPro	506
1330	Db	AATGCAGATTCTACAGTTAAATCTGGGATATCAAAACAGGA-----	1371
507	QY	AlaGlyThrLeuCysLeuArgThrLeuVal-----GluHisSerGlyArgValPhe	523
1372	Db	-----CAGTGTTTACAAACATTGCAAGTCCCAACAGCATCAGAGTGTGTGACC	1422
524	QY	ArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThrIleLeuIle	543
1423	Db	TGTTTACAGTTCAACAGAACTTTGTAATTCACGTCAGATGATGGAACTGTAAACTA	1482
544	QY	TrpAsp 545	
1483	Db	TGGGAC 1488	
RESULT 14			
US-10-245-618-17			
; Sequence 17, Application US/10245618			
; Publication No. US20030143502A1			
; GENERAL INFORMATION:			
; APPLICANT: Reed, Steven			
; APPLICANT: Strohmaier, Helmo			
; APPLICANT: Spruck, Charles			
; APPLICANT: Sangfelt, Olle			
; TITLE OF INVENTION: HDCC4 MODULATES CYCLIN E DEGRADATION			
; FILE REFERENCE: TSRI 79.2			

APPLICANT: Sangreitt, Olie
TITLE OF INVENTION: HCDC4 MODULATES CYCLIN E DEGRADATION
FILE REFERENCE: TSRI 779.2

CURRENT APPLICATION NUMBER: US/10/245,618
 CURRENT FILING DATE: 2002-09-16
 PRIOR APPLICATION NUMBER: US 60/404,116
 PRIOR FILING DATE: 2002-08-15
 PRIOR APPLICATION NUMBER: US 60/322,947
 PRIOR FILING DATE: 2001-09-14
 NUMBER OF SEQ ID NOS: 48
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 17
 LENGTH: 1770
 TYPE: DNA
 ORGANISM: Homo sapien
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1)...(1770)
 S-10-245-618-17

Alignment Scores:
 Seq. No.: 3,04e-70 Length: 1770
 Core: 640.00 Matches: 167
 Percent Similarity: 47.79% Conservative: 92
 Best Local Similarity: 30.81% Mismatches: 211
 Very Match: 21.09% Indels: 72
 Gaps: 15

S-09-601-168b-2 (1-569) x US-10-245-618-17 (1-1770)

14 LysPheMetAsnSerSerGluArgGluAspCysAsnAsnGlyGluProArgLysIle 33
 157 AAGTTGACCATGTTCTGAGTCCGCTCTTTCTTTGGGAAAGAACCATGCAAGTC 216
 34 IleProGluLysAsnSerLeuArgGlnThrTyrAsnSerCysAlaArgLeuAsn 53
 217 -----TCGATATACAGTACCCTCGGCTT----- 243
 54 GlnGluThrValCysLeuAlaSerThrAlaMetLysThrGluAsnCysValAlaLysThr 73
 244 -----GTACCATCTTCAGCA-----ACACCAACAACCTTTGGGACCTC 282
 74 LysLeuAlaAsnGlyThrSerSerMetIleValProLysGlnArgLysLeuSerAlaSer 93
 283 AGACGACCAATGTCGCAAGG-----CAACACGACGCGCAATTCATCT 327
 94 TyrGluLysGluLysGluLeu-----CysValLysTyrPheGluGlnTrpSerGluSer 111
 328 GTCCAGCCACCTACAGGCTCCAGGATGGCTAAATGTTTCAGAGCTGGAGTGACCA 387
 112 AspGlnValGluPheValGluHisLeuIleSerGlnMetCysHisTyrGlnHisGlyHis 131
 388 GAGAAATTCCTTGTCTTAGATGAACCTCATTTGATGTTGTAACCAACAGTAAACAT 447
 132 IleAsnSerTyrLeuLysProMetLeuGlnArgAspPheIleThrAlaLeuProAlaArg 151
 448 ATGATGCAAGTGATAGAACCCCATTTTCACAGAGACTTCATTTCTGCTCCCT----- 501
 152 GlyLeuAspHisIleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAla 171
 502 -----AAAGAGTTGGCACTCTATGTGCTTTTCATCTCTGGAACCAAGACCTGTCTACAA 555
 172 AlaGluLeuValCysLysGluTyrTyrArgValThrSerAspGlyMetLeuTrpLysLys 191
 556 CGAGCTCAGACATCTGCTACTGAGAAATTTGGCTGAGACCAACCTTCTCTGGAGAGAG 615
 192 LeuIleGluArgMetValArgThrAspSerLeuTrpArgGlyLeuAlaGlu----- 208
 616 AAATCAAAAGAA-----GAGGGGATTGATGAACCATTCGCAC 651
 209 -----ArgArgGlyTyrGlyGlnTyrLeuPheLysAsnLysProProAspGlyAsnAla 226
 652 ATCAAGAGAGAGAAA-----GTAAATAAACCCAGGTTTCATACACAGT 693
 227 ProProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleGluThrIle 246

Db 694 CCATGGAAAGTGCATAC-----ATCAGACAG-----CACAGATT 729
 Qy 247 GluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSerGluThr 266
 Db 730 GATACTAACTGGAGCGAGAGAACTCAAAATCTCTAAGGTG---CTGAAGAGCATGAT 786
 Qy 267 SerLysGlyValTyrCysLeuGlnTyrAspGlnLysIleValSerGlyLeuArgAsp 286
 Db 787 GATCATGCTGATCACATGCTTACAGTTTGTGTGTAACCGAATAGTTAGTGTCTGATGAC 846
 Qy 287 AsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHis 306
 Db 847 AACACTTTAAAGAGTTTGGTCACAGCTCAGGCAAAATGCTCTGAGAACAATTAGTGGGCAT 906
 Qy 307 ThrGlySerValLeuCysLeuGlnTyrAspGlnValIleIleThrGlySerSerAsp 326
 Db 907 ACAGTTGGAGTATGGTCATCACAAAATGAGACACATCATCTATTAGTGATCTACAGAT 966
 Qy 327 SerThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIleHis 346
 Db 967 CGGACACTCAAAAGTGTGGAATCGAGAGACTGGAGAATGTATACACACCTTATATGGGCAT 1026
 Qy 347 CysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSerLysAsp 366
 Db 1027 ACTTCCACTGTGGTGTGATGATCTTCTGAAAAGAGTTGTAGCGGTCTTCGAGAT 1086
 Qy 367 ArgSerIleAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgArgValLeu 386
 Db 1087 GCCACTCTTAGGTTTGGGATATTGAGACAGGCCAGCTTTTA-----CATGTTTGT 1137
 Qy 387 ValGlyHisArgAlaAlaValAsnValValAspPheAspLysTyrIleValSerAla 406
 Db 1138 ATGGGTCATGTTGCAGCAGTCCGCTGTTCATATGATGCGAGAGGGTGTGTAGTGA 1197
 Qy 407 SerGlyAspArgThrIleLysValTrpAsnThrSerThrCysGluPheValArgThrLeu 426
 Db 1198 GCATATGATTTTGTAAAGGTGGGATCCAGAGACTGAAACCTGCTACACACGTTG 1257
 Qy 427 AsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAspArgLeuValValSerGly 446
 Db 1258 CAGGGGCATCTAATAGAGTCTATTTCATACAGTTTGTATGATGATCCATGCTGGTGA 1317
 Qy 447 SerSerAspAsnThrIleArgLeuTrpAspIleGluCysGlyValaCysLeuArgValLeu 466
 Db 1318 TCTCTTGATACATCAATCGTGTGGGATGGGAGACAGGGAATTCGATTCACACGTTA 1377
 Qy 467 GluGlyHisGluGluLeuValArgCysIleArgPheAspAsnLysArgIleValSerGly 486
 Db 1378 ACAGGCGACCATGCTGTTAAACAGTGGAAATGGAACCTCAAGACCAATATTCCTGTCTGGG 1437
 Qy 487 AlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAlaLeuAspProArgAlaPro 506
 Db 1438 AATGCAGATCTACAGTTAAATCTGGGATATCAAAACAGGA----- 1479
 Qy 507 AlaGlyThrLeuCysLeuArgThrLeuVal-----GluHisSerGlyArgValPhe 523
 Db 1480 -----CAGTGTTTACAAACATTGCAAGTCCCAACAAAGCATCAGAGTCTGTGACC 1530
 Qy 524 ArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThrIleLeuIle 543
 Db 1531 TGTTTACAGTTCAACAGAACTTTGTAATTACCAGCTCAGATGATGGAACTGTAAACTA 1590
 Qy 544 TrpAsp 545
 Db 1591 TGGGAC 1596

RESULT 15

US-10-245-618-5
 ; Sequence 5, Application US/10245618
 ; Publication No. US20030143582A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven
 ; APPLICANT: Strohmaier, Heimo

APPLICANT: Spruck, Charles
 APPLICANT: Sangfelt, Olie
 TITLE OF INVENTION: HCD4 MODULATES CYCLIN E DEGRADATION
 FILE REFERENCE: TSRI 779.2
 CURRENT APPLICATION NUMBER: US/10/245,618
 CURRENT FILING DATE: 2002-09-16
 PRIOR APPLICATION NUMBER: US 60/404,116
 PRIOR FILING DATE: 2002-08-15
 PRIOR APPLICATION NUMBER: US 60/322,947
 PRIOR FILING DATE: 2001-09-14
 NUMBER OF SEQ ID NOS: 48
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 5

LENGTH: 1884

TYPE: DNA

ORGANISM: Homo sapien

FEATURE:

NAME/KEY: CDS

LOCATION: (1)...(1884)

US-10-245-618-5

Alignment Scores:

Seq. No.:	3,38e-70	Length:	1884
Score:	640.00	Matches:	167
Percent Similarity:	47.79%	Conservative:	92
Best Local Similarity:	30.81%	Mismatches:	211
Query Match:	21.09%	Indels:	72
JB:	12	Gaps:	15

US-09-601-168B-2 (1-569) x US-10-245-618-5 (1-1884)

14	LysPheMetAsnSerSerGluArgGluAspCysLenAnGlyGluProProArgLysLeile	33
271	AAAGTTGGACCATGCTTCGAGCTCGCTCTTTTCTTTGGGAAAGAAACCATGCAAGTC	330
34	IleProGluLysAsnSerLeuArgGlnThrTyrAsnSerCysAlaArgLeuCysLeuAsn	53
331	-----TCGAATATACAGTACCACCTGGGCTT-----	357
54	GlnGluThrValCysLeuAlaSerThrAlaMetLysThrGluAsnCysValAlaLysThr	73
358	-----GTACCATGTTTCAGCA-----ACACCAACAACCTTTTGGGACCTC	396
74	LysLeuAlaAsnGlyThrSerSerMetIleValProLysGlnArgLysLeuSerAlaSer	93
397	AGAGCAGCAATGGCCCAAGG-----CAACCAACGACGCCGAAATACATCT	441
94	TyrGluLysGluLysGluLeu-----CysValLysTyrPheGluGlnTrpSerGluSer	111
442	GTCCAGCCACCTACAGGCTCCAGGAATGGCTAAATATGTTTCAGAGCTGGAGTGACCA	501
112	AspGlnValGluPheValGluHisLeuLeileSerGlnMetCysHisTyrGlnHisGlyHis	131
502	GAGAAATTCCTTCTTTAGTAGAAGTCTCATGATAGTGTGTAACCAACACAGTAAACAT	561
132	IleAsnSerTyrLeuLysProMetLeuGlnArgAspPheIleThrAlaLeuProAlaArg	151
562	ATGATCGAAGTCATAGAACCCAGTTTCAACGAGACTTCATTCATTGCTCCCT-----	615
152	GlyLeuAspHisIleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAla	171
616	-----AAGAGTTGGCACTCTATGTGCTTCATTCCTGGAACCCAAAGACCTCTCAAA	669
172	AlaGluLeuValCysLysGluTrpTyrArgValThrSerAspGlyMetLeuTrpLysLys	191
670	GCAGCTCAGACATGTCCTACTGGAGATTTTGGCTGAAGCAACCTTCTCTGGAGAGAG	729
192	LeuIleGluArgMetValArgThrAspSerLeuTrpArgGlyLeuAlaGlu-----	208
730	AAATGCCAAGAA-----GAGGGGATTGATGACCACTTGCAC	765
209	-----ATGArgGlyTrpGlyGlnTyrLeuPheLysAsnLysProProAspGlyAsnAla	226

766	ATCAAGAGAGAGAAAA-----GTAATAAAACCAGGTTTTCATACACAGT	807
227	ProProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleGluThrIle	246
808	CCATGGAAAAGTGATAC-----ATCAGACAG-----CACAGAATT	843
247	GluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSerGluThr	266
844	GATACTAACTGGAGCGGAGGAGAACTCAATCTCTTAAGGTG-----CTGAAAGGACATGAT	900
267	SerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLeuArgAsp	286
901	GATCATGTGATCATCATGCTTACAGTTTGTGGTAACCGAATAGTTAGTGGTCTCTGATGAC	960
287	AsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHis	306
961	AACACTTTAAAGTTTGGTCAGCAGTCACAGGCAATGTCTGAGAACATTAGTGGGACAT	1020
307	ThrGlySerValLeuCysLeuGlnTyrAspGluArgValIleIleThrGlySerSerAsp	326
1021	ACAGGTGGAGTATGTCATCACAAATGAGAGACAACATCATCATTTAGTGGATCTACAGAT	1080
327	SerThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisHis	346
1081	CGGACATCAAAAGTGTGGAATGCGAGACTGCGGAATGTATACACACCTTATATGGCAT	1140
347	CysGluAlaValLeuHisLeuArgPheAsnAsnGlyMetMetValThrCysSerLysAsp	366
1141	ACTTCCACTGTGCGTGTGATGCATCTTCATGAAAGAGAGTGTAGCGGTCTTCGAGAT	1200
367	ArgSerIleAlaValTrpAspMetAlaSerProThrAspIleThrLeuArgValLeu	386
1201	GCCACTTTCAGGTTTGGGATATTGAGACAGCGCAGTGTTTA-----CATGTTTTG	1251
387	ValGlyHisArgAlaAlaValAsnValValAspPheAspLysTyrIleValSerAla	406
1252	ATGGTTCATGTCAGCAGTCCGCTGTGTTCAATATCATGCGCAGAGGGTGTGTAGTGA	1311
407	SerGlyAspAspThrIleLysValTrpAsnThrSerThrCysGluPheValArgThrLeu	426
1312	GCATATGATTTTATGGTAAAGGTGTGGATCCAGAGACTGAAACCTGTCTACACACGTTG	1371
427	AsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAspArgLeuValValSerGly	446
1372	CAGGGGCATACTAATAGAGTCTATTTCATGATTTGATGTTCCATGTTGGTGGTGA	1431
447	SerSerAspAsnThrIleArgLeuTrpAspIleGluCysGlyAlaCysLeuArgValLeu	466
1432	TCTCTTGATACATCAATCCGTTTGGGATGTGGAGACAGGGAATTGCATTTCACAGTTA	1491
467	GluGlyHisGluGluLeuValArgCysIleArgPheAspAsnLysArgIleValSerGly	486
1492	ACAGGGCACCACTCGTTTAAAGTGGATGGAATCGAACTCAAGACATATTCTTGCTCTGGG	1551
487	AlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAlaLeuAspProArgAlaPro	506
1552	AATGCAGATTCTACAGTTAAATCTGGGATATCAAAACAGGA-----1593	
507	AlaGlyThrLeuCysLeuArgThrLeuVal-----GluHisSerGlyArgValPhe	523
1594	-----CAGTGTTTACAAACATTCGAAGTCCCAAGCTCCAGACATCAGATGCTGTGACC	1644
524	ArgLeuGlnPheAspGluPheGlnIleValSerSerHisAspAspThrIleLeuIle	543
1645	TGTTTACAGTTTCAACAGAACTTTGTAATTACAGCTCAGATGATGAACCTGTAACCTA	1704
544	TrpAsp 545	
1705	TGGGAC 1710	

Search completed: October 22, 2003, 18:30:55

Job time : 414 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

M protein - nucleic search, using frame_plus_p2n model

run on: October 22, 2003, 14:52:38 ; Search time 100 seconds

(without alignments)

2511.470 Million cell updates/sec

Title: US-09-601-168B-2

Perfect score: 3034

Sequence: 1 MDPAEVLQKALKPMNSSE.....PAAQAEPSPFSRTYTVISR 569

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

MODEL=frame_p2n.model -DEV=xlpl
Q=/cgn2_1/USPTO.spool_p/US09601168/runat_22102003_100546_17813/app_query.fasta_1.711
DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
MODE=LOCAL -OUTFMT=ptc -NORW=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
USER=US09601168@cgn2_1_185@runat_22102003_100546_17813 -NCPU=6 -ICPU=3
NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	520	17.1	2481	3	US-08-899-578-1
2	392	12.9	2237	4	US-08-914-999-7
3	365.5	12.0	2085	2	US-08-283-917-8
4	365.5	12.0	2085	2	US-08-961-716-8
5	337	11.1	3465	4	US-08-914-999-5
6	334.5	11.0	1909	4	US-09-620-312D-730
7	308	10.2	2359	1	US-08-188-582-4
8	308	10.2	2359	1	US-08-646-715-4
9	299.5	9.9	7042	3	US-09-092-508-1
10	299.5	9.9	7042	4	US-09-435-115-1
11	299.5	9.9	7042	4	US-09-098-310-1
12	299.5	9.9	7042	4	US-09-690-364-21

13	299.5	9.9	7075	3	US-09-092-508-15	Sequence 15, Appl
14	299.5	9.9	7075	4	US-09-435-115-15	Sequence 15, Appl
15	298.5	9.8	2152	1	US-08-188-582-17	Sequence 17, Appl
16	298.5	9.8	2152	1	US-08-646-715-17	Sequence 17, Appl
17	291	9.6	5152	4	US-09-690-364-10	Sequence 10, Appl
18	287	9.5	3747	4	US-09-690-364-17	Sequence 17, Appl
19	268.5	8.8	1211	3	US-09-063-743-2	Sequence 2, Appl
20	268.5	8.8	1211	4	US-09-590-540-2	Sequence 2, Appl
21	268	8.8	2272	4	US-09-108-857-1	Sequence 1, Appl
22	264.5	8.7	1738	4	US-09-620-312D-615	Sequence 615, App
23	264.5	8.7	2100	4	US-09-013-118-2	Sequence 2, Appl
24	256	8.4	3067	4	US-09-016-434-1285	Sequence 1285, Ap
25	253	8.3	7518	4	US-09-620-312D-1051	Sequence 1051, Ap
26	246.5	8.1	1115	1	US-08-190-802A-19	Sequence 19, Appl
27	246.5	8.1	1115	3	US-08-477-346-19	Sequence 19, Appl
28	246.5	8.1	1115	4	US-08-473-089-19	Sequence 19, Appl
29	246.5	8.1	1115	4	US-08-487-072A-19	Sequence 19, Appl
30	239	7.9	7881	2	US-08-751-189-1	Sequence 1, Appl
31	239	7.9	7881	2	US-09-060-836-1	Sequence 1, Appl
32	239	7.9	7881	3	US-09-184-445-1	Sequence 1, Appl
33	233	7.7	2119	3	US-09-032-372-6	Sequence 6, Appl
34	226	7.4	126	3	US-09-172-841-2	Sequence 2, Appl
35	226	7.4	126	4	US-08-951-621-2	Sequence 2, Appl
36	226	7.4	7886	2	US-08-751-189-2	Sequence 2, Appl
37	226	7.4	7886	2	US-09-060-836-2	Sequence 2, Appl
38	226	7.4	7886	3	US-09-184-445-2	Sequence 2, Appl
39	225.5	7.4	3363	4	US-09-620-312D-222	Sequence 222, App
40	222	7.3	1221	4	US-08-965-600-2	Sequence 2, Appl
41	222	7.3	1221	4	US-09-489-506-2	Sequence 2, Appl
42	220	7.3	2369	4	US-09-302-769-20	Sequence 20, Appl
43	219	7.2	1517	3	US-09-180-783-1	Sequence 1, Appl
44	217	7.2	4911	4	US-09-718-692-1	Sequence 1, Appl
45	217	7.2	4911	4	US-09-718-652-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-899-578-1
; Sequence 1, Application US/08899578
; Patent No. 6087153
; GENERAL INFORMATION:
; APPLICANT: Greenwald, Iva
; APPLICANT: Hubbard, E. Jane
; TITLE OF INVENTION: SEL-10 AND USES THEREOF
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,578
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/53200/JPW/AKC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 278-0525
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2481 base pairs
; TYPE: nucleic acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 91..1854
; JS-08-899-578-1

```

Alignment Scores:

```

Pred. No.: 7,88e-50 Length: 2481
Score: 520.00 Matches: 150
Percent Similarity: 43.85% Conservative: 78
Sest Local Similarity: 28.85% Mismatches: 222
Query Match: 17.14% Indels: 70
DB: 3 Gaps: 17

```

JS-09-601-168B-2 (1-569) x US-08-899-578-1 (1-2481)

```

2y 72 LysThrLysLeuAlaAsnGlyThrSerSerMetIle----- 83
Db 193 GAAAGCTCGTACAGCAATGCTCTTCCAGCTACAATGCTGACAAATTATCGTCCTCC 252
2y 84 valprolysglnarglys-----LeuSerAlaSerTyrGluLysGluLysGluLeu--- 100
Db 253 AGACCTTGGCAACACAACTTGATTTATCGGCTTCTCCCTCTCGAAACAACAGCACTCAAT 312
2y 101 -----CysValLysTyrPheGluGlnTyrSerGluSerAspGlnValGlu 115
Db 313 CCGCGTGTGCAACATTTGATCGCATTTATTCAGGATCTATCAAGCGCGGAACAAATGGAT 372
2y 116 PheValGluHisLeuLeuSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyr 135
Db 373 GCATTCACAGCTGTCTGAGCAATCCAAATGATCAACAAATATTCGACAGTGGGTGCCATT 432
2y 136 LeuLysProMetLeuGlnArgPheIleThrAlaLeuProAlaArgGlyLeuAspHis 155
Db 433 ATTGAACCCCATTTCCAGCGGATTTCTCTCTGCTCCCTGTC-----GAG 480
2y 156 IleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAlaAlaGluLeuVal 175
Db 481 CTGGAATGAAATCTCTCAACATTTAAACGGGATATGACCTGCTCAAAAGTGCACAGGTG 540
2y 176 CysLysGluTyrTyrArgValThrSerAspGlyMetLeuTyrLysLysLeu---IleGlu 194
Db 541 TCGAAAAATTGGAAATTTGATCTGAAATTCGAAATTCGAAATTTGGAAGAGTCTTGCTCGAA 600
2y 195 ArgMetVal-----ArgThrAspSerLeuTyrArgGlyLeuAla 207
Db 601 GAGTTTAAACATCATCCAGATCCACAGACCGAGTTTACTGGTGGCGGCAAGAACTGCA 660
2y 208 GluArgArgGlyTyrGlyGlnTyrLeuPheLysAsnLysProProAspGlyAsnAlaPro 227
Db 661 ATTGCTGTCTGGA-----GTCACATTTCTCTGATCACAATTCAGCCA 699
2y 228 ProAsn---SerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIle----- 243
Db 700 TGTGATCTTAATGTTATCGATCTTAAGTTCAGAGCTTTGGAGATATCTTCGAAGCG 759
2y 244 -----GluThrIleGluSerAsnTyrArgCysGly 253
Db 760 GCTGCTGACAAGTCACGTTATCTTCGAGCGCGGATAAAATTCGAAAGAACTGGAATCGCAAT 819
2y 254 ArgHisSerLeuGlnArgIleHisCysArgSerGluThrSerLysGlyValTyrCysLeu 273
Db 820 CCAATATGGGTCAGCAGTG---CTACAGAGGACAGAGATCATGTAATCACTTGTATG 876
2y 274 GlnTyrAspAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTyrAsp 293
Db 877 CAAATTCATGATGCTGTGGTGAAGTCTGACGATAACACTCTTAAAGATGTTGGTGT 936
2y 294 LysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCysLeu 313
Db 937 ATTGACAAAGGAGAGGTTATGTACACACTAGTCGCGCACACTGGAGGAGTTTGGACATCA 996

```

```

Qy 314 GlnTyrAspGlu-----ArgValIleIleThrGlySerSerAspSerThrValArgVal 331
Db 997 CAGATTTCCTCAATGCGGAAGATATATTGTAGCGGGTCCACTGATAGAACTGTAAAGTT 1056
Qy 332 TrpAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisHisCysGluAlaValLeu 351
Db 1057 TGGAGTACTGTAGATGGTTCACTTCTTCATACACTCAAGGACATACCTCCACTGTTGGA 1116
Qy 352 HisLeuArgPheAsnAsnGlyMetMetValThrCysSerLysAspArgSerIleAlaVal 371
Db 1117 TGCATGGCTATGCTGCTTCCATATCTGCTCACTGGATCACGAGATACCACTTCTCGTGA 1176
Qy 372 TrpAspMetAlaSerProThrAspIleThrLeuArgArgValLeuValGlyHisArgAla 391
Db 1177 TGGAGCTAGAAATCCGAGCTCACCTGCA-----ACTTTACATGGCCATCATGCA 1227
Qy 392 AlaValAsnValValAspPheAspAspLysTyrIleValSerAlaSerGlyAspArgThr 411
Db 1228 GCGTTCGATGCGTTCAATTCGATCGAAACAACGTGTTTCGGGAGGATATGATTTTACC 1287
Qy 412 IleLysValTyrAsnThrSerThrCysGluPheValArgThrLeuAsnGlyHisLysArg 431
Db 1288 GTTAAATTTGGAATGCTCATACTCGGAGATGTATCCGCTACTCTGACCGGTCAACAAT 1347
Qy 432 GlyIleAlaCysLeuGlnTyrArgAspArg-----LeuValValSerGlySerSerAsp 449
Db 1348 AGAGTTTATCTCTCTCTTCTTGAAGCGGCGATCGTGTGCTCTGCTCTCTGCGAC 1407
Qy 450 AsnThrIleArgLeuTyrAspIle-----GluCysGlyAlaCysLeuArgValLeu 466
Db 1408 ACTTCAATTCGGGTGGGATTTTACACGACCGGAAGCGCAAGATGTGTGGCTCTTTTG 1467
Qy 467 GluGlyHisGluGluLeuValArgCysIleArgPheAspAsnLysArgIleValSerGly 486
Db 1468 CAAAGACACACTCACTTACATCCGGAATGCACCTTCGAGGCAATATCTCGTGTATGC 1527
Qy 487 AlaTyrAspGlyLysIleLysValTyrAspLeuValAlaAlaLeuAspProArgAlaPro 506
Db 1528 AATGCAGATAGCATGTTAGATATGGATAT-----CAC 1563
Qy 507 AlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGlyArgValPheArgLeuGln 526
Db 1564 GAGGGAAC---TGTGTACACATGCTTCTGGACATCGATCCGCTATCACTTCTTCA 1620
Qy 527 ---PheAspGluPheGlnIleValSerSerSerHisAspAspThrIleLeuIleTyrAsp 545
Db 1621 TGGTTTGGACGAAATATGTTAGTACGAGTAGTACGATGGAACGTGCAAAATTTGGGAT 1680

```

RESULT 2

```

US-08-914-999-7
; Sequence 7, Application US/08914999
; Patent No. 6346406
; GENERAL INFORMATION:
; APPLICANT: RYAZANOV, Alexey G.
; APPLICANT: Hait, William N.
; APPLICANT: Pavur, Karen S.
; TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)
; TITLE OF INVENTION: AND METHODS OF USE THEREFOR
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

[illegible][illegible]

```

365 365 1813 385 1864 405 1924 425 1984 445 2044 465 2104 483 2164
Db Db QY Db QY Db QY Db QY Db QY Db QY Db QY Db QY Db QY
1813 AATGATAAAACTCAAGGTTTGGGATTGAAACT-----TTTCGTTGTAACCTAC 1863
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
385 ValLeuValGluHisArgAlaAlaValAsnValValAspPheAspAspLysTyrIleVal 404
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1864 ACTCTAAAGGTCATCTAAATGGGTCCACCCTATCTGTATATTAGGTACCAATCTCTAC 1923
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
405 SerAlaSerGlyAspArgThrIleLysValTTPAsnThrSerThrCysGluPheValArg 424
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1924 AGTGGCTCCTATGATAAACTATAGAGTTTGGGATTTAAAGAGTTTAGAATGTTCCGCT 1983
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
425 ThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyrArgAspArgLeuValVal 444
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1984 ACTTTAAGAGGCCATCATAGATGGGTTCGACATATAGTGAATTTGTGATAAATTTATTTT 2043
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
445 SerCysSerSerAspAsnThrIleArgLeuTTPAspIleGluCysGlyAlaCysLeuArg 464
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2044 ACTCCTAGTCACCATATACAAATTAAATTTTCGGATTTAGAAACATTAAGATGTAATACA 2103
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
465 ValLeuGluGlyHisGluGluLeuValArgCysIleArgPhe-----AspAsnLysArg 482
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2104 ACTTTGAAGGACATATGCAACCGTTCAATGTTAGCAGTTTGGGAAGATAAAAAATGT 2163
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
483 IleValSerGlyAlaTyrAspGlyLysIleLysValTTP 495
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2164 GTTATTATTGTTAGTCATGATCAATCAAGTATTAGAGTTTGG 2202
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 3
US-08-283-917-8
; Sequence 8, Application US/08283917
; Patent No. 5849557
; GENERAL INFORMATION:
; APPLICANT: ADACHI, HIDEKI
; APPLICANT: TSUJIMOTO, MASAFUMI
; APPLICANT: INOUE, KEIZO
; APPLICANT: ARAI, HIROYUKI
; TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME
; TITLE OF INVENTION: AND GENE THEREOF
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:

```

```

RESULT 3
US-08-283-917-8
; Sequence 8, Application US/08283917
; Patent No. 5849557
; GENERAL INFORMATION:
; APPLICANT: ADACHI, HIDEKI
; APPLICANT: TSUJIMOTO, MASAFUMI
; APPLICANT: INOUE, KEIZO
; APPLICANT: ARAI, HIROYUKI
; TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME
; TITLE OF INVENTION: AND GENE THEREOF
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
;

```

ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER &
 ADDRESS: NEUSTADT, P.C.
 STREET: 1755 S. Jefferson Davis Highway, Suite 400
 CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/283,917
 FILING DATE: 03-AUG-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 209943/1993
 FILING DATE: 03-AUG-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Oblon, No. 5849557man F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 2292-030-0
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 413-3000
 TELEFAX: (703) 413-2220
 TELEX: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2085 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: cDNA
 ORIGINAL SOURCE:
 ORGANISM: Bos taurus
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 844..2073
 US-08-283-917-8

Alignment Scores:
 Pred. No.: 4,83e-32 Length: 2085
 Score: 365.50 Matches: 140
 Percent Similarity: 38.83% Conservative: 98
 Best Local Similarity: 22.84% Mismatches: 210
 Query Match: 12.05% Indels: 165
 DB: 2 Gaps: 22
 US-09-601-168b-2 (1-569) x US-08-283-917-8 (1-2085)
 QY 20 GluArgGluAspCysAsnAsnGlyGluProProArgLysIlelleProGluLysAsnSer 39
 DB 463 GAGCGCGCTGAC-----CCCAACCAACCATCCGAGCATCCACCCCAA 507
 QY 40 LeuArgGlnThrTyrrAsnSerCysAlaArgLeuCysValLeuAsnGln-----GluThr 56
 DB 508 ATCCGGCAGGANTTTCTTTTCGCGCTCGCTCTTCAACGGAGCTGCCTTTTGACGTT 567
 QY 57 ValCysLeuAlaSerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAla 76
 DB 568 GTAACACTGAGCTTCGAGGCCCTCAGCCATCTCTCTCGAATCTCCCACTCGTATAGGA 627
 QY 77 AsnGlyThrSerSerMetIleValProLysGlnArgLysLeuSerAlaSerTyrrGluLys 96
 DB 628 AACGAGTGCCTCGCTTA-----ACCTCCCAAGTG 657
 QY 97 GluLysGluLeuCysValLysTyrrPheGluGlnTrpSerGluSerAspGlnValGluPhe 116
 DB 658 GAATGAACCTTACTTTGTAATATCTC----- 684
 QY 117 ValGluHisLeuIleSerGlnMetCysHisTyrrGlnHisGlyHis----- 131

DB 685 -----CTGGTTACACGTTGGATTCACTTGTGAAAGAAATCATTTTCCCTCGCGTAA 735
 QY 132 -----IleAsnSerTyrrLeuLysProMet 139
 DB 736 AGCCACTTAGTGGCTTATTAAATTATAAATCCAGGGGTTGCAAGCTTTTGG-----ATT 789
 QY 140 LeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsn 159
 DB 790 TTCAGAGGAGGACATACCACTATA-----TCGAATAGCTTGTATATTACAGCAAAATG 846
 QY 160 IleLeuSer----- 162
 DB 847 GTGCTGTCCCAAGACACAGAGATGAATAATCGAGCTATAGCAGATTATCTTCGTCA 906
 QY 163 -----TyrrLeuAspAlaLysSerLeuCysAlaAlaGlu----- 173
 DB 907 AATGGCTACGAAGAAGCATATTTCAGTTTAAAAAGGAAGCTGAATTAGATATGAATGAA 966
 QY 174 -----LeuValCysLysGluTrpTyrrArgValThrSerAsp 185
 DB 967 GAATTAGATAGAAATATGCTGGTCTTTTGGAAAAAATGG-----ACATCTGTT 1017
 QY 186 GlyMetLeuTrpLysLysLeuIleGluArgMetValArgThrAspSerLeuTrpArgGly 205
 DB 1018 ATTAGATTACAAAAGAAGGTTATGGAATTAGAAATCAAAAGTTAAATGAAGCAAAAGAA 1077
 QY 206 LeuAlaGluArgAspGlyTyrrGlyGlnTyrrLeuPheLysAsnLysProProAspGly--- 224
 DB 1078 TTACGTCGGGTGGACCTCTTGGTCAG-----AAAAGAGACCCAAAGAATGGATT 1128
 QY 225 AsnAlaProProAsnSerPheTyrrArgAlaLeuTyrrProLysIlelleGlnAspIleGlu 244
 DB 1129 CCCCCTCCACAGAAATAATATGCATTGAGTGGTCAT----- 1164
 QY 245 ThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSer 264
 DB 1165 -----AGGAGT 1170
 QY 265 GluThrSerLysGlyValTyrrCysLeuGlnTyrrAspAspGlnLysIlelleValSerGlyLeu 284
 DB 1171 CCAGTCACTCGAGTCATTTTCCATCTCTGTGTC-----AGTGTATGTGCTCTGCTTCA 1224
 QY 285 ArgAspAsnThrIleLysIleTrpAspLysAsnThrLeuGluCysLysArgIleLeuThr 304
 DB 1225 GAGGATGCTACAAATAAGGTGGGATTATGAGACTGGAGATTTTGACCAACTTTAAG 1284
 QY 305 GlyHisThrGlySerValLeuCysLeuGlnTyrrAspGlu-----ArgValIleIleThr 322
 DB 1285 GGGCATACAGACTCTGTACAGGATATTTCATTCGACCACAGTGGCAAGCTTCTGGCTTCA 1344
 QY 323 GlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeuAsnThr 342
 DB 1345 TGTTCTGCAGATATGACCATTAAGCTATGGGATTTTCAGGCGCTTTGTAATGCATCAGAAC 1404
 QY 343 LeuIleHisHisCysGluAlaValLeuHisLeuArgPhe-----AsnAsnGlyMetMet 360
 DB 1405 ATGCATGCCCATGACCACAAATGTTCTTCAGTACCATCATGCCCATGGAGATCATATA 1464
 QY 361 ValThrCysSerLysAspArgSerIleAlaValTrpAspMetAlaSerProThrAspIle 380
 DB 1465 GTGTCCTGCTCAAGGATATAAACTATAAAATGTGGAAAGTGCAAACTGGCTACTGTGTG 1524
 QY 381 ThrLeuArgValLeuValGlyHisArgAlaAlaValAsnValVal-----AspPhe 398
 DB 1525 -----AAGACATTCACAGGACACAGAGATGGGTAGCTATGTTGGTGGCGCCAAATCAA 1575
 QY 399 AspAspLysTyrrIleValSerAlaSerGlyAspArgThrIleLysValTrpAsnThrSer 418
 DB 1576 GACGSCACTCTGATAGCCAGCTGTTCCCAATGACCAGACTGTGCGTGTATGGTGTAGCA 1635
 QY 419 ThrCysGluPheValArgThrLeuAsnGlyHisTyrrArgGlyIleAlaCysLeuGlnTyrr 438
 DB 1636 ACAAGGAATCAAGGCTGAGCTTCGAGAACATGCATGTGGTAGAATGCATTTCTCTGG 1695

QY 439 ArgAspArg----- 441
Db 1696 GTCCTGAAAGCTCATATCTCTCCATCTCTGAAGCAACAGGATCTGAGACTAAAAAGT 1755
QY 442 -----LeuValValSerGlySerSerAspAsnThrIleArgLeuTrpAsp 456
Db 1756 GGCACACCTGGGCATCTTACTGTCCGATCCAGGCAAGACTATCAAGATGGGAT 1815
QY 457 IleGluCysGlyAlaCysLeuArgValLeuGluGlyHisGluGluLeuValArgCysIle 476
Db 1816 GTCAGTACTGGCATGTCCTTATGACCCCTGGTGGTCATGATACTGGGTAGTGGAGTT 1875
QY 477 ArgPheAspAsn-----LysArgIleValSerGlyAlaTyAspGlyLysIleVal 494
Db 1876 CTGTTTCATCTGGGGGAAGTTATTTTGGTGGCTGATGACAAAGACCCCTGGCGTG 1935
QY 495 TrpAspLeuValAlaAlaLeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThr 514
Db 1936 TGGGATTAC-----AAGAACAAGCGATGCATGAAGACC 1968
QY 515 LeuValGluHisSerGlyArgValPheArgLeuGlnPheAspGluPhe-----GlnIle 532
Db 1969 CTAATCGCATGAACACTTGTATCTCTCTCTGATTCTCATAAGACGCGCCCATATGTG 2028
QY 533 ValSerSerSerHisAspAspThrIleLeuIleTrpAsp 545
Db 2029 GTTACTGGCAGTGTAGATCAACAGTAAGGTGTGGAG 2067

RESULT 4

US-08-961-716-8
Sequence 8, Application US/08961716
Patent No. 5880272
GENERAL INFORMATION:
APPLICANT: ADACHI, HIDEKI
APPLICANT: TSUJIMOTO, MASAFUMI
APPLICANT: INOUE, KEIZO
APPLICANT: ARAI, HIROYUKI
TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME
TITLE OF INVENTION: AND GENE THEREOF
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &
ADDRESSEE: NEUSTADT, P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,716
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/283,917
FILING DATE: 03-AUG-1994
APPLICATION NUMBER: JP 209943/1993
FILING DATE: 03-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5880272man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 2292-030-0
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:

LENGTH: 2085 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
ORIGINAL SOURCE: Bos taurus
ORGANISM: Bos taurus
FEATURE:
NAME/KEY: CDS
LOCATION: 844..2073
US-08-961-716-8
Alignment Scores:
Pred. No.: 4,83e-32 Length: 2085
Score: 365.50 Matches: 140
Percent Similarity: 38.83% Conservative: 98
Best Local Similarity: 22.84% Mismatches: 210
Query Match: 12.05% Indels: 165
DB: 22 Gaps: 22
US-09-601-168B-2 (1-569) x US-08-961-716-8 (1-2085)
QY 20 GluArgGluAspCysAsnAsnGlyGluProProArgLysIleIleProGluLysAsnSer 39
Db 463 GAGCGGCGCTGAC-----CCACCGAACCATCCGACGATCCACCCACCAA 507
QY 40 LeuArgGlnThrTyAsnSerCysAlaArgLeuCysLeuAsnGln-----GluThr 56
Db 508 ATCCGGCAGGATTTCTTTCTCGCGTCGGCTCTTCAACGGGAGCTGCTTTTGACGTT 567
QY 57 ValCysLeuAlaSerThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAla 76
Db 568 GTAACACTGAGCTTCGAGGCCCTCAGCCATTCTCTCGAATCTCCCATCTCGTATAGA 627
QY 77 AsnGlyThrSerSerMetIleValProLysGlnArgLysLeuSerAlaSerTyArgLys 96
Db 628 AACGCAGTGCCTGCCTTA-----ACCTCCAGGTG 657
QY 97 GluLysGluLeuCysValLysTyPheGluGlnTrpSerGluSerAspGlnValGluPhe 116
Db 658 GAATGAACCTTACTTGTGATATCTC----- 684
QY 117 ValGluHisLeuIleSerGlnMetCysHisTyArgHisGlyHis----- 131
Db 685 -----CTGTTACAGTGGATTCATTGTGAAGAATCAITTTCCCTCGCTGAA 735
QY 132 -----IleAsnSerTyLeuLysProMet 139
Db 736 AGCCACTTAGTGGCTTATTAAATTAATAATCCAGGGGTTCGAAAGCTTTTG-----ATT 789
QY 140 LeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGluAsn 159
Db 790 TTCAGAGAGGGGACATAACCACTATA---TCGAATAAGCTTGATATTACAGCCAAATG 846
QY 160 IleLeuSer----- 162
Db 847 GTGCTGTCCCGACAGACACAGATGAACCTAAATCGAGCTATAGCAGATTATCTCTGTTCA 906
QY 163 -----TyrlleuAspAlalysSerLeuCysAlaAlaGlu----- 173
Db 907 AATGGCTTACGAGAGACATATTCAGTTTTTAAAGAGGAGCTGAATTAGATATGAATGAA 966
QY 174 -----LeuValCysLysGluTrpTyArgValThrSerAsp 185
Db 967 GAATTAGATAAGAAATATGCTGCTCTTTTGGAAAAAATGG-----ACATCTGTT 1017
QY 186 GlyMetLeuTrpLysLeuIleGluArgMetValArgThrAspSerLeuTrpArgGly 205
Db 1018 ATTAGATTACAAAGAGGTTATGGAATTAGAATCAAAAGTTAAATGAAGCAAGCAAGAA 1077
QY 206 LeuAlaGluArgGlyTrpGlyGlnTyLeuPheLysAsnLysProProGly--- 224
Db 1078 TTACGTGGGTGGACCTCTTGTGTAG-----AAAGAGACCCAAAGAAATGGATT 1128

```

QY 225 AsnAlaProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleGlu 244
Db 1129 CCCGGTCCACCAAGAAATATGATTCAGTGGTCAT-----1164
QY 245 ThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSer 264
Db 1165 -----AGGAGT 1170
QY 265 GluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLysIleValSerGlyLeu 284
Db 1171 CCAGTCACTCGAGTCATTTCCATCTGTGTC-----AGTGTAGGTCCTCGTCTCA 1224
QY 285 ArgAspAsnThrIleLysIleTyrAspLysAsnThrLeuGluCysLysArgIleLeuThr 304
Db 1225 GAGGATGCTACAATTAAGGTGTGGGATATGAGATTTTGAACGAACCTCTTAAG 1284
QY 305 GlyHisThrGlySerValLeuCysLeuGlnTyrAspGlu-----ArgValIleIleThr 322
Db 1285 GGGCATAACAGACTGTGACAGGATATTTCAATCGACACAGTGGCAAGCTTCTGGCTCA 1344
QY 323 GlySerSerAspSerThrValArgValTyrAspValAsnThrGlyGluMetLeuAsnThr 342
Db 1345 TGTCTGCAGATATGACCATTAAGCATGCGATTTTCAAGGCTTTGAATGCATCAGAAC 1404
QY 343 LeuIleHisCysGluAlaValLeuHisLeuArgPhe-----AsnAsnGlyMetMet 360
Db 1405 ATGATGGCCATGACCAACAATGTTCTTCAGTACCCATCATGCCCAATGGAGATCATATA 1464
QY 361 ValThrCysSerLysAspArgSerIleAlaValTyrAspMetAlaSerProThrAspIle 380
Db 1465 GTGTCTCCCTCAGGGATAAACTATAAAATGTGGAGTGCAGAACTGCTACTGTGTG 1524
QY 381 ThrLeuArgArgValLeuValGlyHisArgAlaAlaValAsnValVal-----AspPhe 398
Db 1525 -----AAGCATTTCAGGACACAGAGATGGTACGTATGTGGCGCCAAATCAA 1575
QY 399 AspAspLysTyrIleValSerAlaSerGlyAspArgThrIleLysValTyrAsnThrSer 418
Db 1576 GACGGCACTGTAGTACGAGCTGTTCGAATGACAGACTGTGGGTGTATGGGTCTGTAGCA 1635
QY 419 ThrCysGluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCysLeuGlnTyr 438
Db 1636 ACAAGGAATGCAGGCTGAGCTTCGGAACATGAGCATGTGTAGATTCATTTCTCTGG 1695
QY 439 ArgAspArg-----441
Db 1696 GCTCCTGAAAGCTCATATTTCTTCATCTCTGAAGCAACAGGATCTGAGACTAAAAAAGT 1755
QY 442 -----LeuValValSerGlySerSerAspAsnThrIleArgLeuTyrAsp 456
Db 1756 GGCAAACCTGGGCATTTCTTACTGTCCGGATCCAGGACAGAGACTATCAAGATGGGAT 1815
QY 457 IleGluCysGlyAlaCysLeuArgValLeuGluGlyHisGluLeuValArgCysIle 476
Db 1816 GTCAGTACTGGATGTCCTTATGACCTGTGGTGCATGATCACTGGGTACGTGAGTT 1875
QY 477 ArgPheAspAsn-----LysArgIleValSerGlyAlaTyrAspGlyLysIleLysVal 494
Db 1876 CTGTCTCATTTCTGGGGGAAAGTTATTTTGTAGTTGCGCTGATCACAAGACCCCTGCGCGTG 1935
QY 495 TrpAspLeuValAlaAlaLeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThr 514
Db 1936 TGGGATTAAC-----AAGAACAGCGATGATGAAGACC 1968
QY 515 LeuValGluHisSerGlyArgValPheArgLeuGlnPheAspGluPhe-----GlnIle 532
Db 1969 CTCAATGCGATGAACACTTTGTACCTCTTGGATTTCCATBAGAGCGGCCCATATGTG 2028
QY 533 ValSerSerHisAspAspThrIleLeuIleTyrAsp 545
Db 2029 GTTACTGGCAGTGTAGATCAACAGATAAGGTGTGGAG 2067

```

```

RESULT 5
US-08-914-999-5
; Sequence 5, Application US/08914999
; Patent No. 6346406
; GENERAL INFORMATION:
; APPLICANT: Ryazanov, Alexey G.
; APPLICANT: Hait, William N.
; APPLICANT: Pavur, Karen S.
; TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)
; TITLE OF INVENTION: AND METHODS OF USE THEREFOR
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/914,999
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 601-1-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3465 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Dictyostelium discoideum
; US-08-914-999-5

Alignment Scores:
Pred. No.: 2,24e-28 Length: 3465
Score: 337.00 Matches: 79
Percent Similarity: 50.00% Conservative: 52
Best Local Similarity: 30.15% Mismatches: 103
Query Match: 11.11% Indels: 28
DB: Gaps: 7

US-09-601-168B-2 (1-569) x US-08-914-999-5 (1-3465)
QY 246 IleGluSerAsnTrpArgCys-----GlyArgHisSerLeuGlnArg 259
Db 2692 GTCAATGACAAATTTGGAATGTTTATACACTGTCTCAATGGTCATAGAAAATCAATTTGAAAGT 2751
QY 260 IleHisCysArgSerGluThrSerLysGlyValTyrCysLeuGlnTyrAspAspGlnLys 279
Db 2752 ATCGCTGTGTAATAGT-----AATTAC 2772
QY 280 IleValSerGlyLeuArgAspAsnThrIleLys-----IleTrpAspLysAsnThr 296
Db 2773 ATTTTCACTTCATCCCTGATACACCATCAAGTTTCATATCATTCGTAGTGTAAACACC 2832
QY 297 LeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCysLeuGlnTyrAsp 316
Db 2833 ---AAATGATATAGACACATTTGGTTGGTCACACTGGTGAAGTTAATTGTGTCTGCGCCAAT 2889

```

```

Qy 317 GluArgValIleThrGlySerSerAspSerThrValArgValTrpAspValAsnThr 336
Db 2890 GAAAAATATCTTTTCAGTTGTAGTTATGATAAAACTATCAAGGTTTGGGATTTGCAACC 2949
Qy 337 GlyGluMetLeuAsnThrLeuIleHis---HisCysGluAlaValLeuHisLeuArgPhe 355
Db 2950 TTTAAAGAAATTAATAATTTGAGGGGTTCATCAAAAGTACATTAAACATTTGGCTTTG 3009
Qy 356 AsnAsnGlyMetMetValThrCysSerLysAspArgSerIleAlaValTrpAspMetAla 375
Db 3010 AGTCGACGTTATCTTTTAGTGGTGGTAACGATCAATCATTTACGTTTGGGATACTGAA 3069
Qy 376 SerProThrAspIleThrLeuArgArgValLeuValGlyHisArgAlaValAsnVal 395
Db 3070 ACA-----CTTAGTATGCTTTTCAATATGCAAGGTCATGAAGATTTGGGTACTCTCT 3120
Qy 396 ValAspPheAspLysIleValSerLysAlaSerGlyAspArgThrIleLysValTrp 415
Db 3121 CTTCATTGTCACCGTAGTTATCTTTCTCAACTCAAAAGATTAATGTCATCAAGATTGG 3180
Qy 416 AsnThrSerThrCysGluPheValArgThrLeuAsnGlyHisLysArgGlyIleAlaCys 435
Db 3181 GATCTCTCAAAATTTTCAGTTGTATGATACCTCTAAAGGTCATTGGAATTTCTGCTCAAGT 3240
Qy 436 LeuGlnTyArgAspArgLeuValValSerGlySerSerAspAsnThrIleArgLeuTrp 455
Db 3241 TGTGTCGTAAAGATCGTTATCTATACAGTGGTTCGAAGATAAATCAATCAAAAGTTTGG 3300
Qy 456 AspIleArgCysGlyAlaCysLeuArgValLeu---GluGlyHisGluGluLeuValArg 474
Db 3301 GATCTCGATACACTTGAATGTGTTCACACCATTCCAAATCTCATTTCTTTGGGTGTAAAA 3360
Qy 475 CysIleArgPheAspAsnLysArgIleValSerGlyAlaTyArgSerGlyLysIleLysVal 494
Db 3361 TGTTTAAAGTTTTCAAATAATCAATTCATTCGCTTTCGCTTTCGATGTTCAATTAAAGTT 3420
Qy 495 TrpAsp 496
Db 3421 TGGGAA 3426

```

RESULT 6

US-09-620-312D-730

Sequence 730, Application US/09620312D

Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyan

APPLICANT: Chen, Rui-hong

APPLICANT: Zhao, Qing A.

APPLICANT: Wehrman, Tom

APPLICANT: Xue, Aidong J.

APPLICANT: Yang, Yonghong

APPLICANT: Wang, Jian-Rui

APPLICANT: Zhou, Ping

APPLICANT: Ma, Yunqing

APPLICANT: Wang, Zhiwei

APPLICANT: John Tillinghast

APPLICANT: Drmanac, Radoje T.

TITLE OF INVENTION: No. 6569662el Nucleic Acids and

FILE REFERENCE: 784CIP2B

CURRENT APPLICATION NUMBER: US/09/620,312D

PRIOR FILING DATE: 2000-07-19

PRIOR FILING DATE: 2000-07-19

PRIOR FILING DATE: 2000-04-25

PRIOR FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 1105

SOFTWARE: pt_FL_genes Version 1.0

```

; SEQ ID NO 730
; LENGTH: 1909
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (147)..(1511)
US-09-620-312D-730

Alignment Scores:
Pred. No.: 1,648-28 Length: 1909
Score: 334.50 Matches: 110
Percent Similarity: 37.50% Conservative: 67
Best Local Similarity: 23.31% Mismatches: 144
Query Match: 11.03% Indels: 151
DB: Gaps: 14

US-09-601-168B-2 (1-569) x US-09-620-312D-730 (1-1909)

Qy 103 LysTyrPheGluGlnTrpSerGluSerAspGlnValGluPhe----- 116
Db 156 AAGGACTTTTGAGACATGGCTTGATAACATTTCTGTACATTTCTTCTGACGGACTTG 215
Qy 117 -----ValGluHisLeuIleSerGlnMetCysHisTyrGlnHisGlyHis 131
Db 216 CAGAAAAAATGAAACTCTGGATCACCTGATTAGTCTGAGTGGGGCAGTCCAGCTCAGGCAT 275
Qy 132 IleAsnSerTyrLeuLysProMetLeuGlnArgAspPheIleThrAlaLeuProAlaArg 151
Db 276 CTCTCCAATAAAGCTAGAGACTCTCTCAAGGGGAGTCTCTCAAACTCCTTCCCTG--- 332
Qy 152 GlyLeuAspHisIleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCysAla 171
Db 333 -----GAGCTCAGTTTTTATTGTTAAATGGCTCGATCCTCAGACTTTTACTCACA 383
Qy 172 AlaGluLeuValCysLysGluTrpTyrArgValThrSerAspGlyMetLeuTrpLysLys 191
Db 384 TGCTGCTCGTCTCTAAACAGTGAATGAAGTGATAAGT----- 422
Qy 192 LeuIleGluArgMetValArgThrAspSerLeuTrpArgGlyLeuAlaGluArgGly 211
Db 423 -----GCCTGTACAGAGGTGGCAGACTGCATGTGTAATAAATTTGGGC 464
Qy 212 TrpGlyGlnTyLeuPheLysAsnLysProProAspGlyAsnAlaProProAsnSerPhe 231
Db 465 TGG-----CAGATAGATGATTCTGTTTCAGGACGCTTTGGCAG 500
Qy 232 TyrArgAlaLeuTyProLysIleIleGlnAspIleGluThrIleGluSerAsnTrpArg 251
Db 501 TGGAGAAGGTTTATTGGAAGGCTATTTGAGAATGAAGCAACTGGAGGAC----- 551
Qy 252 CysGlyArgHisSerLeuGlnArgIleHisCysArgSerGluThrSerLysGlyValTyr 271
Db 552 -----CATGAAGCCTTTGAAACCTCGTCA----- 575
Qy 272 CysLeuGlnTyArgAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIle 291
Db 575 ----- 575
Qy 292 TrpAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeu 311
Db 576 -----TTAATTGGACACAGTCCAGAGTGTAT 602
Qy 312 CysLeuGlnTyArgGluArgValIleIleThrGlySerSerAspSerThrValArgVal 331
Db 603 GCACCTTTACTACAAAGATGGACTTCTCTGACAGGTCAGATGACTTCTGTCGCAAGCTG 662
Qy 332 TrpAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisCysGluAlaValLeu 351
Db 663 TGGGATGTGACGACAGGGCAGTGGCTTTATGGCATCCAGACCCAC----- 707
Qy 352 HisLeuArgPheAsnAsnGlyMetMetValThrCysSerLysAspArgSerIleAlaVal 371
Db -----

```


708	----	ACTGTGCA	716
372	TpAspMetAlaSerProThrAspIleThrLeuArgValLeuValGlyHisArgAla	391	
716	----	716	
392	AlaValAsnValValAspPheAspAspLysTyrIleValSerAlaSerGlyAspArgThr	411	
717	----	767	
412	IleLysValTpcAsnThrSerThrCysgluPheValArgThrLeuAsnGlyHisLysArg	431	
768	GTGGCTGTCTGGGAATGGAGTTCGGAGCCAGGACCCAGCACATTCGGGGGCACACGGGG	827	
432	GlyIleAlaCysLeuGlnTyrArgAspArgLeu-----ValValSerGlySerSerAsp	449	
828	CGCGTATTAGCGTGGGACTACAATGATGAACATGGATATCTTGGTGACGGCGTCTGCAGAC	887	
450	AsnThrIleArgLeuTpcAspIleGlyCysGlyAlaCysLeuArgValLeuGluGlyHis	469	
888	TTCACGTGAAAGTATGGGCTTATCTGCTGGGACATGCCGTGAACACACTCAOOGGGCAC	947	
470	GluGlu-----LeuValArgCys-----IleArgPhe	478	
948	ACGGAATGGGTCAACAAAGTAGTTTTGCAGAGTGCAAAGTCAAGTCTCTTTCGCACAGT	1007	
479	AspAsnLysArgIleValSerGlyAlaTyrAspGlyLysIleLysValThrAspLeuVal	498	
1008	CCTGGAGACTACATCCCTCTTAAGTGCAGACAAATATGAGATTAGATTTCGGCAATGGG	1067	
499	AlaAlaLeuAspProArgAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHis	518	
1068	AGCAAAATCAACTGTAG-----TGCTTAAGACATTTGCTGTCTCT	1109	
519	SerGlyArgValPhe-----ArgLeuGlnPheAspGluPheGlnIleValSer	534	
1110	GAGGATAGAAGTATCTGCCTGCACCCAGACTTCATTTTGATGGCAATACATATGCTGT	1169	
535	SerSerHisAspAspThrIleLeuIleThrAspPhe	546	
1170	AGTTACACACTTGT-----CTTACACAGTGGGACTTT	1202	

RESULT 7

US-08-188-592-4
; Sequence 4, Application US/08188582
; Patent No. 553410
; GENERAL INFORMATION:
; APPLICANT: Ttian, Robert
; APPLICANT: Comai, Lucio
; APPLICANT: Dynlact, Brian D.
; APPLICANT: Hoey, Timothy
; APPLICANT: Ruppert, Siegfried
; APPLICANT: Tanese, Naoko
; APPLICANT: Wang, Edith
; APPLICANT: Weinzierl, Robert O.J.
; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOEBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188,582
; FILING DATE: 28-JAN-1994

Db 1086 AAAG-----CTTAAGGCTCTCAGGAAGCCAGCAGCGTTTAGCCCTCAGCAAGGATCA 1139
 Qy 266 -----ThrsrLysGlyValTyrCysLeuGl 274
 Db 1140 ACTGCCCTCTGCCGCTTCTACAGCGTGCTTAATCCCATCAGGGGCTAACCTGTGCCGA 1199
 Qy 274 nTyrAspAspGlnLys-----IleValSerGlyLeuArgAspAsnThrIleLysIleTr 292
 Db 1200 GATTTCAGACGATTCCACGATGTTGGCTGTGGATTTGGCGATTCAGGTGAGGATTG 1259
 Qy 292 pAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCy 312
 Db 1260 GTCAATGACGCCGCGAAGCTGCTAGCTGAAGATGCAGATTCC-----1305
 Qy 312 sleuGlnTyrAspGluArgValIleIleThrGlySerSerAspSerThrValArgValTr 332
 Db 1306 -----CTTCGCGAAGCTGGACAGGAATCGCGGATATCAATGTGCGTATGCT 1352
 Qy 332 pAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisCysGluAlaValLeuHi 352
 Db 1353 GGATGACCAAGTGTGAGGTAAACAGGAGCTTAATGGGTACACCGGACCGGTATACCG 1412
 Qy 352 sleuArgPhe-----AsnAsnGlyMetMetValThrCysSerLysAspArgSerIleAl 370
 Db 1413 CTGTGCTTTCGCCCGAGATGAACCTGTCTCATGTTCCGAGGACAGCACCATAG 1472
 Qy 370 aValTrpAspMetAlaSerProThrAspIleThrLeuArgArgValLeuValGlyHisAr 390
 Db 1473 GCTGTGGTCTCTGCTCACTGCTGCTGGTGTAGTCACTACCGC-----GGGCACGT 1523
 Qy 390 gAlaAlaValAsnValValAspPhe-----AspAspLysTyrIleValSerAlaSerGl 408
 Db 1524 TTACCGGTGTGGATGTTCGCTTTCGCCGCGATGCTATTTTGTTCCTGTGCTGA 1583
 Qy 408 YAspArgThrIleLysValTrpAsnThrSerThrCysGluPheValArgThrLeuAsnGl 428
 Db 1584 CGACAAACTGCTGCTGTGGGCGACGATTCCAATCAAGCGTGTGGCTATTCGTGGG 1643
 Qy 428 YHisLysArgGlyIleAlaCysLeuGlnTyr-----ArgAspArgLeuValSerGl 446
 Db 1644 TCACCTGTGCGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1703
 Qy 446 YSerSerAspAsnThrIleArgLeuTrpAspIleGluCysGlyAlaCysLeuArgValLe 466
 Db 1704 TTCTAGCGATCGACGAGTACGCTGTGGCAACATGACCGGTGCTGCTGCTGCTGCT 1763
 Qy 466 uGluClyHisGluGluLeuValArgCysIleArgPheAsp-----AsnLysArgIleVa 484
 Db 1764 GACGGGCGACAGGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1823
 Qy 484 LserGlyAlaTyrAspGlyLysIleLysValTrpAspLeuValAlaAlaLeuAspProAr 504
 Db 1824 CTCGGGTTTCAGTAGATCAATATCATCATCTGGGATCTG-----1863
 Qy 504 gAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGlyArgValPheAr 524
 Db 1864 ----TCGAACGAGTCCCTG--GTACACCCCTGTTGAGGACACACTAGCAGTCTGTGACC 1916
 Qy 524 gleuGlnPhe-----AspGluPheGlnIleValSerSerHisAspAspThrIleLe 542
 Db 1917 GATCACCCTTACTCGCGATGGACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1976
 Y 542 uIleTrpAspPhe 546
 b 1977 TCTGTGGGACTTT 1989

RESULT 8
 S-08-646-715-4
 Sequence 4, Application US/08646715
 Patent No. 5637686
 GENERAL INFORMATION:
 APPLICANT: Tjian, Robert
 APPLICANT: Comai, Lucio

APPLICANT: Dynlact, Brian D.
 APPLICANT: Hoey, Timothy
 APPLICANT: Ruppert, Siegfried
 APPLICANT: Tanese, Naoko
 APPLICANT: Wang, Edith
 APPLICANT: Weinzierl, Robert O.J.
 TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
 NUCLEIC ACIDS ENCODING TAPS AND METHODS OF USE
 NUMBER OF SEQUENCES: 36
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FLEHR, HOEBACH, TEST, ALBRITTON & HERBERT
 STREET: 4 Embarcadero Center, Suite 3400
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94111-4187
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/646,715
 FILING DATE: 09-MAY-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/188,582
 FILING DATE: 28-JAN-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Osman, Richard A.
 REGISTRATION NUMBER: 36,627
 REFERENCE/DOCKET NUMBER: A-57650-2/AUT/RAO
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 781-1989
 TELEFAX: (415) 398-3249
 TELEX: 910 277299
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2359 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 NAME/KEY: CDS
 LOCATION: 49..2160
 US-08-646-715-4
 Alignment Scores:
 Pred. No.: 2,74e-25 Length: 2359
 Score: 308.00 Matches: 121
 Percent Similarity: 40.95% Conservative: 94
 Best Local Similarity: 23.05% Mismatches: 193
 Query Match: 10.15% Indels: 117
 DB: 1 Gaps: 22
 US-09-601-168B-2 (1-569) x US-08-646-715-4 (1-2359)
 Qy 107 GlnTrpSerGluSerAsp-GlnValGluPheValGluHisLeuIleSerGlnMetCys-- 125
 Db 513 CAGTGAGCTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 563
 Qy 126 -----HisTyrGlnHisGlyHisIleAsnSerTyrLeu-----LysPro----- 138
 Db 564 TCTCGAGGCTACTACATAGAGGGTCTTTTCAACCTTTTGTGCTGTCTAAGCCCGAGGA 623
 Qy 139 -MetLeuGlnArgAspPheIleThrAlaLeuPro-----Al 150
 Db 624 GCTGCTGGAGATGACCTGCTAGTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 683
 Qy 150 aArgGlyLeuAspHisIleAlaGluAsnIleLeuSerTyrLeuAspAlaLysSerLeuCy 170
 Db 684 CAGGGACTCGACTCTCTGTTCAGGCGACACATTCAG-----GATCGCGCGAGGAAGT 737

QY 170 salalagluLeuValCysLysGluTrpTyArgValThrSerAspGlyMetLeuTrpLy 190
Db 738 GGTGCAGATATGTTTCCAAAGTACTTTCGACATACAGAGGCGATGGCGGCAA 797
QY 190 slyLeuileGluArgMetValArgThrAspSerLeuTrpArgGlyLeuAlaGluArgAr 210
Db 798 CAAGCTG-----CAGTGGCTCGCCACCGCGGCTCGCACCTCGAGAGGCCAAGGACA 851
QY 210 gGlyTrpGlyGlnTyLeuPhe----- 217
Db 852 CGACAACAAAATGCGGGTGTACTACGACTCTCAAGGAGGTGGACTTTCAGACTCTGAC 911
QY 217 ----- 217
Db 912 CACTCCAGCGCGGCACCAGAGGAGGAGCGATGATCCGGATCGCCGGATCGTCGAA 971
QY 218 -LysAsnLysPro-----ProAspGlyAsnAl 226
Db 972 AAAGAAAAAGCCAAAAAAGGATCCCTGCTGCGAAAAAGTCCAAAGTCGGATCCGAATGC 1031
QY 226 aProProAsnSerPheTyArgAlaLeuTyProLysIleileGlnAspIleGluThrIl 246
Db 1032 TCCA-----TCCATCGACAGAAATCCCTCGCGGAAGTTCGAGCAAGTTGCT 1085
QY 246 eGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSerGlu-- 265
Db 1086 AAAG-----CTTAAGGCTCTCAGGGAAGCCAGCAAGCGTTTAGCCCTCAGCAAGGATCA 1139
QY 266 -----ThrSerLysGlyValTyCysLeuGl 274
Db 1140 ACTGCCCTCTGCCGCTTCTACCGTCTTAAATTCCTCATCAGGCGGTAACTGTGCCGA 1199
QY 274 nTyArgAspGlnLys-----IleValSerGlyLeuArgAspAsnThrIleLysIleTr 292
Db 1200 GATTTCACAGCAGTCCACAGATGTGGCTGTGGATTGGCGATTCAGCGAGAGATTG 1259
QY 292 pAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCy 312
Db 1260 GTCAATTGACGCCCGGAGTGCCTGACGAGTGAAGGATCGAGATTCC----- 1305
QY 312 sLeuGlnTyArgGluArgValIleThrGlySerSerAspSerThrValArgValTr 332
Db 1306 -----CTTCGCGAAGTCCGACCAAGAAATCGCGCGATATCAATGTGCGTATGCT 1352
QY 332 pAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisCysGluAlaValLeuHi 352
Db 1353 GGATGACCGAAGTGTGAGGTAAACAGGAGCTTAATGGGTACACCGGACCGGTATACCG 1412
QY 352 sLeuArgPhe-----AsnAsnGlyMetMetValThrCysSerLysAspArgSerIleAl 370
Db 1413 CTGTGCTTTGCCCGCGAGATGAACCTGTGTCTCATGTTCCGAGGACAGCACCAATAAG 1472
QY 370 aValTrpAspMetAlaSerProThrAspIleThrLeuArgArgValLeuValGlyHisAr 390
Db 1473 GCTGTGGTCTCTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1523
QY 390 gAlaAlaValAsnValAlaValPhe-----AspAspLysTyIleValSerAlaSerGl 408
Db 1524 TTACCGGCTGTGGGATGTTGCTTTGCGCGCGATGGCTACTATTGTTCTGTGTGTA 1583
QY 408 yAspArgThrIleLysValTrpAsnThrSerThrCysGluPheValArgThrLeuAsnGl 428
Db 1584 CGACAAAATGCTGCTGCTGTGGGCCACCGAATTCCAATCAAGCGTTTGCCTGTTGCTGG 1643
QY 428 yHisLysArgGlyIleAlaCysLeuGlnTy-----ArgAspArgLeuValValSerGl 446
Db 1644 TCACCTGTGCGAGCGGTGATGTGTACAAATTCATCCCAATTCCTCAATATGATGGCCACCG 1703
QY 446 ySerSerAspAsnThrIleArgLeuTrpAspIleGluCysGlyAlaCysLeuArgValLe 466
Db 1704 TTCTAGCGATCGCACGGTACGCTGTGGGACACACATGACCGGTCTAGTCTGCTACGCTGAT 1763

QY 466 uGluGlyHisGluGluLeuValArgCysIleArgPheAsp-----AsnLysArgIleVa 484
Db 1764 GAGGCGCCACAGGATCGGTGAGTTCTCTGGCTTCTCCGCTCGCGCGGTATCTGCGC 1823
QY 484 lSerGlyAlaTyArgGlyLysIleLysValTrpAspLeuValAlaLeuAspProAr 504
Db 1824 CTCGGGTTTCAGTAGATCAATATCATCATCTGGGATCTG----- 1863
QY 504 gAlaProAlaGlyThrLeuCysLeuArgThrLeuValGluHisSerGlyArgValPheAr 524
Db 1864 ----TCGAACGATCCCTG---GTCAACACCTGTGTAGGCACATAGCAGTGTGACCA 1916
QY 524 gLeuGlnPhe-----AspGluPheGlnIleValSerSerHisAspAspThrIleLe 542
Db 1917 GATCACCITTAGTCGATCGATGGAACAGTCTGTGCTGCGACGCGCTTGGATAACAATCTAAC 1976
QY 542 uIleTrpAspPhe 546
Db 1977 TCTGTGGGACTTT 1989
RESULT 9
US-09-092-508-1
; Sequence 1, Application US/09092508
; Patent No. 6291643
; GENERAL INFORMATION:
; APPLICANT: Henzel, William J.
; TITLE OF INVENTION: APAF-1, AN ACTIVATOR OF C ASPASE-3
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Weiter & Schmidt
; STREET: 3100 No. 6291643west Center, 90 South Seventh St
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/092,508
; FILING DATE: 05-JUN-1998
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 60/048,807
; FILING DATE: 05-JUN-1997
; APPLICATION NUMBER: 60/055,258
; FILING DATE: 07-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Ph.D., Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 11669.6USU1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7042 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 578...4159
; OTHER INFORMATION:
; US-09-092-508-1
Alignment Scores: 1 59e-23 Length: 7042
Pred. No.: 299.50 Matches: 154
Score:

Percent Similarity:	33.77%	Conservative:	102
Best Local Similarity:	20.32%	Mismatches:	195
Query Match:	9.87%	Indels:	308
DB:	3	Gaps:	34
US-09-601-168B-2 (1-569) x US-09-092-508-1 (1-7042)			
QY	2	AspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGluArg	21
DB	1769	GAAGAAGTTGAAGACATACTCGAGAG-----TTGTAAATAAGTCTCTTTA	1816
QY	22	GluAspCysAsnGlyGluProArg-----LysIle	33
DB	1817	TTCTGTGATCGGAATGGAAAGTCGTTTGTATTATTACATGATCTTCAAGTAGATTIT	1876
QY	34	IleProGluLysAsnSerLeuArg-----	41
DB	1877	CTTACAGAGAAGATTGACGCCAGCTTCAGGATCTACATAAGAGATAATCACTCAGTTT	1936
QY	42	GlnThrTyrAsnSerCysAlaArgLeuCysLeuAenGlnGlu-ThrValCysLeuAla	61
DB	1937	CAGAGATATCACCAGCCGATACCTCTTTCACCAGATCAGGAAGACTGTATGTATGGTAC	1996
QY	61	rThrAlaMetLysThrGluAenCysValAlaLysThrLysLeuAlaAsnGlyThrSerSe	81
DB	1997	AACTTT-CTGGCCTATCAC-----	2014
QY	81	rMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeuCy	101
DB	2015	-----ATGCCAGTCCCAAGATGCACAGAGACTTTG	2046
QY	101	sValLysTyrPheGlu-----GlnTyrSerGluSerAspGlnValGluPheVal	118
DB	2047	TGCTTTAATGTTTCCCTGGATTGGAATAAAGCA---AAAACAGAACTTGTAGCCCTGC	2103
QY	118	uHisLeuIleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysPr	138
DB	2104	TCATCTGATTCATGAATTGTGGATACAGACAT-----AT	2139
QY	138	oMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGl	158
DB	2140	ACTAGATGAAAGGATTTGTGCA-----GTCAGTGA	2169
QY	158	uAsnIleLeuSerTyrLeuAspAlaLys-----	167
DB	2170	GAAATTTTCAGAGTTTATCTTTAAATGGACACTTCTGGACGACAGCCATTCCTAA	2229
QY	168	-----SerLeuCysAlaAlaGluLeuValCysLysGluTyrTyr	180
DB	2230	TATTGTACAACTGGGTCTCTGTGACCCGAA-----ACTTCAGAAAGTTTATCAGCAAGC	2283
QY	181	-----ArgValThrSerAspGlyMetLeuTyr-----Ly	190
DB	2284	TAAGCTGCAGGCCAAGCAGGAGTGCATTAAGTATGCTTTACCTGGAATGGATAAACAA	2343
QY	190	sLysLeuIleGluArgMetValArg-----ThrAspSerLeuTyrPar	204
DB	2344	AAAAAACATCACGAATCTTTCCCGCTTAACTGTGCGCCGCCACACAGATGCTGT	2398
QY	204	GGlyLeuAlaGluArgGlyTyrGlyGlnTyrLeuPheLysAsnLysProAspGl	224
DB	2398	-----	2398
QY	224	yAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleGl	244
DB	2399	-----TACCATGCTGCTTTCT-----GAGGATGTC	2427
QY	244	urThrIleGluSerAsnTyrArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSe	264
DB	2428	GAGAATAGCTTCT-----TGCGAGCTGTATAAACCTTACAGGTGTC---AAAGC	2475
QY	264	rgLThrSerLys-----GlyValTyrCysLeuGl	274
DB	3553	AGTGAAGACTTTAGACTCTTGAAAAATTCAGAGCTGCTTCTTGGTTCATTTGATGAAC	3612
DB	2476	TGAACAGGAGAGAAACTTCTAGAAATCAAGGCTCATGAGGATGAAGTGTGTGTGTC	2535
QY	274	nTyr-----AspAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTy	292
DB	2536	ATTCTCTACAGATGACAGATTATTACCAACTGCTCAGTGGATAAAAAAGTGAAGATTG	2595
QY	292	pAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCy	312
DB	2596	GAATTCCTATGACTGGGAACTAGTACACACCTTATGATGAGCACTCAGAGCAAGTCAATTG	2655
QY	312	sLeuGlnTyrAspGluArg-----ValIleIleThrGlySerSerAspSerTh	328
DB	2656	CTGCCATTTCCACCAACAGTAGTCATCTCTCTAGCCACTGGTCAAGTACTGCTT	2715
QY	328	rValArgValTyrAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisCysGl	348
DB	2716	CCTCAAACTTTGGATTGGAATCAAAAAGAAATGTCGAATACCATGTTTGGTGCATACAA	2775
QY	348	uAlaValLeuHisLeuArgPhe-----AsnAsnGlyMetMetValThrCysSerLysAs	366
DB	2776	TTTCACTCAATCACTGCAGATTTCACAGATGATGAAGCTTTTGGCTAGTGTTCAGCTGA	2835
QY	366	pArgSerIleAlaValTyrAspMetAlaSerProThrAsp-----IleThrLeuAr	383
DB	2836	TGGAACCTTAAAGCTTTGGGATGCGACATCAGCAAAATGAGAGGAAAAGCATTAATGTGA	2895
QY	383	g-----	383
DB	2896	ACAGTTCTTCTAAATTTGGAGGACCTTCAAGAGATATGGAAGTAGTAGTGAAGTGTG	2955
QY	384	-----ArgValLeuVal-----	387
DB	2956	TTCTGTGCTGTGATGTGTGAGGATATATGTTGGCAGCAAAAAATAAAATCTTTTGTG	3015
QY	388	-----GlyHisArgAlaAlaValAsnVa	395
DB	3016	GAATACAGACTCAGTTCAAAGGTGGCTGATTCGAGAGACATTTAAGTTGGTTCATGG	3075
QY	395	lValAspPhe-----AspAspLysTyrIleValSerAlaSerGlyAspArgThrIleLy	413
DB	3076	TGTGATGTTTCTCTGATGGATCATCTTTTGCATCTCTGATGACCAACATCAG	3135
QY	413	sValTyrAsnThr-----	417
DB	3136	GCTCTGGAGACAAAGAAATGTAAGAACTCTGCTGTAATGTTAAAGCAAGAGTAGA	3195
QY	417	-----	417
DB	3196	TGTTGTGTTTCAAGAAAATGAAGTATGTTGCTGCTGCTGCTGATGACCATATAAGACGCTGCA	3255
QY	418	-----SerThrCy	420
DB	3256	ACTCATTATGGAAGAACAGGTCAAGATTGATTATCTGACTGAAGCTCAAGTTAGCTGCTG	3315
QY	420	s-----GluPh	422
DB	3316	TTGCTTAAGTCCACATCTTCAGTACATTCATTGGAGATGAAATGGAGCCATTGAGAT	3375
QY	422	eValArgThrLeuAsnGly-----HisLysArgGlyIleAl	434
DB	3376	TTTAGAACTGTAAACAATAGAACTCTTCCAGTCCAGGTTTCAGCAACAAGAACTGATG	3435
QY	434	aCysLeuGlnTyrArg-----AspArgLeuValValSerGlySerSerAspAsnThrIl	452
DB	3436	GCATATCCAGTTCACAGCCGATGAGAGACTCTTTATTTCAAGTTCTGATGCTGAAAT	3495
QY	452	eArgLeuTyrAspIleGluCysGlyAlaCysLeuArgValLeuGluGlyHisGluGluLe	472
DB	3496	TCAGGTATGAATGGCAATTTGGACAAATGTATC---TTTCTACAGGCGCCATCAGGAAC	3552
QY	472	uValArgCysIleArg-----PheAspAsnLysArgIleValSerGlyAlaTyrAspGlyLy	491

```

QY 491 stlelysvlTTPaspLeuValAlaLeuAaspProArgAlaProAlaGlyThrLeuCy 511
Db 3613 AGTGAAGGTAGTAATATTACTTGGAAATAAGAA----- 3649
QY 511 sleuArgThrLeuValGluHisSerGlyArgValPheArg-----LeuGlnPheAspG1 529
Db 3650 ---AAAGACTTTGCTGTCACACAGGTACAGTACTTTCTTGACATTCTCAGCATGC 3705
QY 529 upheGlnlleValSerSerHisAspAspThrleLeulleTrpAspPhe 546
Db 3706 TACCAAGTTTTCATCTCTGCTGCAAGAGCTGCAAGAGTCTGGAGTTT 3757

RESULT 10
US-09-435-115-1
; Sequence 1, Application US/09435115
; Patent No. 6346607
; GENERAL INFORMATION:
; APPLICANT: Henzel, William J.
; TITLE OF INVENTION: APAF-1, AN ACTIVATOR OF C ASPASE-3
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. 6346607 West Center, 90 South Seventh St
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/435,115
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/092,508
; FILING DATE:
; APPLICATION NUMBER: 60/055,258
; FILING DATE: 07-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Ph.D., Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 11669.6USU1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7042 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 578...4159
; OTHER INFORMATION:
US-09-435-115-1

Alignment Scores:
Pred. No.: 1.59e-23 Length: 7042
Score: 299.50 Matches: 154
Percent Similarity: 33.77% Conservative: 102
Best Local Similarity: 20.32% Mismatches: 195
Query Match: 9.87% Indels: 308
DB: 4 Gaps: 34

US-09-601-168b-2 (1-569) x US-09-435-115-1 (1-7042)

```

```

QY 2 AspProAlaGluAlaValLeuGlnGluLysPheMetAsnSerSerGluArg 21
Db 1769 GAAGAGTTGAAGACATACTGCGAGG-----TTTGTAATAAGTCTCTTTTA 1816
QY 22 GluAspCysAsnAsnGlyGluProArg-----Lysile 33
Db 1817 TTCTGTGATCGGAATGGAAGTCGTTTGGTTATTATTATCATGATCTTCAAGTAGATTT 1876
QY 34 IleProGluLysAsnSerLeuArg----- 41
Db 1877 CTTACAGAGAAGAAATGCGAGCGCTTCAGGATCTACATAAGAAAGATATCACTCAGTTT 1936
QY 42 GlnThrTyrAsnSerCysAlaArgLeuGluCysLeuAsnGlnGlu-ThrValCysLeuAlaSe 61
Db 1937 CAGAGATATCACCAGCCGCTACTCTTTCCACAGATCAGGAAGACTGTATGTTGTTAC 1996
QY 61 rThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSerSe 81
Db 1997 AACTTT-CTGGCCTATCAC----- 2014
QY 81 rMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeuCy 101
Db 2015 -----ATGCCAGTCCAGAGATGCACAGGAACTTTG 2046
QY 101 sValLysTyrPheGlu--GlnTrpSerGluSerAspGlnValGluPheVal-----G1 118
Db 2047 TGCTTTAATGTTTTCCTCGATTGGAATAAGCA---AAAAACAGAACTTGTAGGCCCTGC 2103
QY 118 uHisLeulleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysPr 138
Db 2104 TCATCTGATTCAATGATTTGCGAATACAGCAT-----AT 2139
QY 138 oMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaG1 158
Db 2140 ACTAGATGAAGAGGATTGTGCA-----GTCAGTGA 2169
QY 158 uAsnIleLeuSerTyrLeuAspAlaLys----- 167
Db 2170 GAATTTTCAGGAGTTTATTCTTTAAATGGACACCTTTTGACGACAGCCATTTCTTAA 2229
QY 168 -----SerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr----- 180
Db 2230 TATTGTACAACTGGGTCCTCTGTCAGCCGAA-----ACTTCAGAGTTTATCAGACAGC 2283
QY 181 -----ArgValThrSerAspGlyMetLeuTrp-----Ly 190
Db 2284 TAAGCTGCAGGCCAAGCAGGAGGTCGATAATGGAATGCTTTACCTGGAATGATAAACAA 2343
QY 190 sLysLeuIleGluArgMetValArg-----ThrAspSerLeuTipAr 204
Db 2344 AAAAAACATCAGAACTCTTCCGCTTAGTTCTCGCCCCCACCACAGATGCTGTT----- 2398
QY 204 gGlyLeuAlaGluArgArgGlyTrpGlyGlnTyrLeuPheLysAsnLysProProAspG1 224
Db 2398 ----- 2398
QY 224 yAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleG1 244
Db 2399 -----TACATGCTGCTTTTCT-----GAGGATGGTCA 2427
QY 244 uThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSe 264
Db 2428 GAGAATAGCTTCT-----TGTGGAGCTGATAAAACCTTACAGGTGTC---AAAGC 2475
QY 264 rGluThrSerLys-----GlyValTyrCysLeuG1 274
Db 2476 TGAACACAGGAGAGAACTTCTAGAAATCAAGGCTCATGAGGATGAAGTCTTTGTTGTC 2535
QY 274 nTyr-----AspAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTr 292
Db 2536 ATTCTCTACAGATGACAGATTTATAGCAACCTGCTGCTGAGTGAATAAAAGTGAAGATTG 2595
QY 292 pAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCy 312

```

```

2596 GAATCTTATGACTGGGAAGTCTGACACCTATGATGAGCACTCAGACCAAGTCAATG 2655
QY sLeuGlnTyrAspGluArg-----ValIleIleThrGlySerSerAspSerTh 328
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2656 CTGCCATTTCACCAACAGTAGTCATCTCTCTTAGCCACTGGCTCAAGTGAAGTCTT 2715
QY rValArgValTyrAspValLeuThrGlyGluMetLeuAsnThrLeuIleHisCysGl 348
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2716 CCTCAAACTTTGGGATTTGAATCAAAAAGATGTCGAATACCATTTTGGTCATACAA 2775
QY uAlaValLeuHisLeuArgPhe-----AsnAsnGlyMetMetValThrCysSerLysAs 366
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2776 TTCAGTCAATCACTGCAGATTTTCCACGATGATAAGCTTTTGGCTAGTTGTTGACGTGA 2835
QY pArgSerIleAlaValTyrAspMetAlaSerProThrAsp-----IleThrLeuArg 383
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2836 TGAACCTTAAAGCTTTGGGATGCGACATCAGCAAAATGAGAGGAAAGCATTAATGTGAA 2895
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
383 g----- 383
2896 ACAGTCTCTCTAAATTTGGAGGCCCTCAAGAGGATATGGAAGTATAGTAGTGAAGTTG 2955
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
384 -----ArgValLeuVal----- 387
2956 TTCGTGGTCTGCTGATGTTGTCAGGATTAATGGTGAGCAAAATAAATAATCTTTTGTG 3015
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
388 -----GlyHisArgAlaAlaValAsnVa 395
3016 GAATACAGACTCAGCTTCAAAAGTGGCTGATGTCAGAGGACATTTAAGTTGGTTCATGG 3075
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
395 lValAspPhe-----AspAspLysTyrIleValSerAlaSerGlyAspArgThrIleLy 413
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
3076 TGTGATGTTTCTCTGTGATGATCATCTTTTGACATCTCTGATGACACAGACATCAG 3135
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
413 eValTyrAsnThr----- 417
3136 GCTCTGGGACACAAGAAAGTATGTAAAGACTCTGCTGTAATGTTAAAGCAAGATAGA 3195
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
417 ----- 417
3196 TGTGTGTTTCAAGAAAAATGAAGTGATGTCCTTCAGTTGACCATATAGAGCTCTGCA 3255
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
418 -----SerThrCys 420
3256 ACTCATTAATGGAAGACAGGTCAGATTGATTATCTGACTGAAGCTCAAGTTAGCTGTG 3315
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
420 g-----GluPhe 422
3316 TTGCTTAAAGTCCACATCTTCAGTACATTTGGAGATGAAAAATGGAGCCCATTCAGAT 3375
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
422 eValArgThrLeuAsnGly-----HisLysArgGlyIleAl 434
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
3376 TTAGAACTTGTAAACAATAGAAATCTCCAGTCCAGGTTTCAGCACAAGAAACTGTATG 3435
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
434 aCysLeuGlnTyrArg-----AspArgLeuValValSerGlySerSerAspAsnThrI 452
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
3436 GCACATCCAGTTCACAGCCGATGAGAACACTCTTATTCAAGTTCTGATGATGCTGAAAT 3495
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
452 eArgLeuTyrAspIleGluCysGlyAlaCysLeuArgValLeuGluGluHisGluGluLe 472
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
3496 TCAGTATGGAATGGCAATTTGGCAAAATGTATC-----TTTCTAGAGGCCATCAGAAAC 3552
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
472 uValArgCysIleArg-----PheAspAsnLysArgIleValSerGlyAlaTyrAspGly 491
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
3553 AGTGAAGAGCTTTAGACTCTTGAAAAATTCAGAACTGCTTTCTTGGTCAATTTGATGAAAC 3612
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
491 sIleLysValTyrAspLeuValAlaAlaLeuAspProArgAlaProAlaGlyThrLeuCy 511
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
3613 AGTGAAGGTATGGAAATATTATTACTTGGAAATAAAGAA-----LeuGlnPheAspGI 529
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
511 sLeuArgThrLeuValGluHisSerGlyArgValPheArg-----LeuGlnPheAspGI 529
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

3650 -----AAAGACTTTGTCTGTACCAGGTTACAGTACTTCTTGTGACATTTCTCAGATGC 3705
QY 529 uPheGlnIleValSerSerHisAspAspThrIleLeuIleTyrAspPhe 546
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
3706 TACCAAGTTTTCATCTCTGTGACAGACTGCAAGATCTGGAGTTT 3757
RESULT 11
US-09-098-310-1
; Sequence 1, Application US/09098310
; Patent No. 6403765
; GENERAL INFORMATION:
; APPLICANT: Alnemeti, Emad S.
; TITLE OF INVENTION: TRUNCATED APAF-1 AND METHODS OF USE
; FILE REFERENCE: 480140.438
; CURRENT APPLICATION NUMBER: US/09/098.310
; CURRENT FILING DATE: 1998-06-16
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 7042
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (578)...(4159)
US-09-098-310-1
Alignment Scores:
Pred. No.: 1,59e-23 Length: 7042
Score: 299.50 Matches: 154
Percent Similarity: 33.72% Conservative: 102
Best Local Similarity: 20.32% Mismatches: 195
Query Match: 9.87% Indels: 308
DB: 4 Gaps: 34
US-09-601-168B-2 (1-569) x US-09-098-310-1 (1-7042)
QY 2 AspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGluArg 21
Db 1769 GAAGAAGTTGAAGACATATCGAGGAG-----TTGTAAATAGATCTCTTTTA 1816
QY 22 GluAspCysAsnAsnGlyGluProArg-----LysIle 33
Db 1817 TTCTGTGATCGGAATGGAAGTCGTTCTGTTATTATTATCATGATCTTCAAGTAGATTT 1876
QY 34 IleProGluLysAsnSerLeuArg----- 41
Db 1877 CTACAGAGAAATTTGACCCAGCTTCAGATCTACATAGAGATTAATCACTCAGTTT 1936
QY 42 GlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGlu-ThrValCysLeuAlaSe 61
Db 1937 CAGAGATATCACCAGCCGCTACTCTTTCCAGCATATCAGGAAGACTGATGATTGGTAC 1996
QY 61 rThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSerSe 81
Db 1997 AACTTT-CTGGCCTATCAC----- 2014
QY 81 rMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeuCy 101
Db 2015 -----ATGCCAGTCCAGATGCAAGAACTTTG 2046
QY 101 sValLysTyrPheGlu-----GlnTyrSerGluSerAspGlnValGluPheVal-----G1 118
Db 2047 TGCTTTAATGTTTTCCTCGATGGATTGAAACAA-----AAAACAGAACTTTGAGGCCCTGC 2103
QY 118 uHisLeuIleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysPr 138
Db 2104 TCATCTGATTCAATTTGTTGAATACAGAT-----AT 2139
QY 138 oMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaGl 158
Db 2140 ACTAGATGAAAGGATTGTGCA-----GTCAGTGA 2169

```

```
QY 158 uAsnIleLeuSerTyrIleuAspAlaLys----- 167
DB 2170 GAATTTTCAGGAGTTTATCTTTAAATGGACACCTCTCTGGAGCAGACCCATTCCTAA 2229
QY 168 -----SerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr----- 180
DB 2230 TATTGTACAACACTGGGTCTCTGTGAGCCGAA-----ACTTCAGAAGTTTATCAGCAAGC 2283
QY 181 -----ArgValThrSerAspGlyMetLeuTrp-----Lys 190
DB 2284 TAAGCTGAGCCCAAGCAGGAGGTGCGATAATGGAAATGCTTTTACCTGGAATGGATAACAA 2343
QY 190 sLysLeuIleGluArgMetValArg-----ThrAspSerLeuTrpAr 204
DB 2344 AAAAAACATCACGAATCTTTCCCGCTTAGTTGTCGCGCCCAACACACAGATGCTGTT----- 2398
QY 204 gGlyLeuAlaGluArgGlyTrpGlyGlnTyrLeuPheLysAsnLysProAspG1 224
DB 2398 ----- 2398
QY 224 yAsnAlaProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleG1 244
DB 2399 -----TACCATGCTGCTTTTCT-----GAGATGGTCA 2427
QY 244 uThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgse 264
DB 2428 GAGAAATAGCTTCT-----TGTGAGCTGATAAAACCTTACAGGTGTC-----AAAGC 2475
QY 264 rGluThrSerLys-----GlyValTyrCysLeuG1 274
DB 2476 TGAACAGGAGAGAAACTTCTAGAAATCAAGGCTCATGAGAGTGAAGTGTGTTGTTGTC 2535
QY 274 nTyr-----AspAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTr 292
DB 2536 ATTCTCTACAGATGACAGATTATAGCAACCTGCTCAGTGCATGATAAAAGTGAAGATTG 2595
QY 292 pAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCy 312
DB 2596 GAATTTCTATGACTGGGAACCTAGTACACACCTATGATGAGCACTCAGAGCAAGTCAATTG 2655
QY 312 sLeuGlnTyrAspGluArg-----ValIleIleThrGlySerSerAspSerth 328
DB 2656 CTGCATTTCCACACAGTAGTCATCTCTCTTAGCCACTGGGTCAAGTCAAGTCTCTT 2715
QY 328 rValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisCysG1 348
DB 2716 CCTCAAACTTTGGGATTTGAATCAAAAGAAATGCGAAATACCATGTTTGGTCAACAAA 2775
QY 348 uAlaValLeuHisLeuArgPhe-----AsnAsnGlyMetMetValThrCysSerLysAs 366
DB 2776 TTCAGTCAATCAGTCGAGATTTTCCAGAGATGATPAAGCTTTTGGCTAGTTGTTCACTGA 2835
QY 366 pArgSerIleAlaValTrpAspMetAlaSerProThrAsp-----IleThrLeuAr 383
DB 2836 TGGAACTTAAAGCTTTGGGATGCGACATCAGCAATGAGAGGAAAGCAATTAATGGA 2895
QY 383 g----- 383
DB 2896 ACAGTTCTCTCTAAATTTGGAGGACCTCAAGAGGATATGGAAGTATAGTGAAGTTG 2955
QY 384 -----ArgValLeuVal----- 387
DB 2956 TTCGTGGTCTGCTGATGTCAGAGATATGTTGGCGCAAAAATAAAATCTTTTGTG 3015
QY 388 -----GlyHisArgAlaAlaValAsnVa 395
DB 3016 GAATACAGACTCAGCTTCAAAAGGTGGTGTATGTCAGAGGACATTTAAGTTGGTTCTATG 3075
QY 395 lValAspPhe-----AspAspLysTyrIleValSerAlaSerGlyAspArgThrIleY 413
DB 3076 TGTGATGTTTCTCTGATGATCATCATTTTTCACATCTTCTGATGACCAACAATCAG 3135
```

```
QY 413 sValTrpAsnThr----- 417
DB 3136 CTTCTGGGAGCAAGAAAGATATGTAAGAACTCTGCTGTGTAATGTTAAAGCAAGAGTAGA 3195
QY 417 ----- 417
DB 3196 TGTGTGTTTCAAGAAAATGAAGTGAATGGTCTTGGCAGTTGACCATATAAAGCGTCTGCA 3255
QY 418 -----SerThrCy 420
DB 3256 ACTCATTAATGGAAGAACAGGTGAGATTGATTATCTGACTGAAGCTCAAGTTAGCTGCTG 3315
QY 420 s-----GluPh 422
DB 3316 TTGCTTAAGTCCACATCTTCTAGTACATTCATTTGGAGATGAAATGGAGCCATTGAGAT 3375
QY 422 eValArgThrLeuAsnGly-----HisLysArgGlyIleAl 434
DB 3376 TTTAGAACTTGTAAACAATAGAATCTTCCAGTCCAGGTTTTCAGCACAAAGAAACTGTATG 3435
QY 434 aCysLeuGlnTyrArg-----AspArgLeuValValSerGlySerSerAspAsnThr11 452
DB 3436 GCATCCAGTTTCACGCCGATGGAAGACTCTTATTTCAAGTTCTGATGATGCTGAAAT 3495
QY 452 eArgLeuTrpAspIleGluCysGlyAlaCysLeuArgValLeuGluGlyHisGluLe 472
DB 3496 TCAGGTATGAATTTGGCAATTTGGCAAAATGTATC---TTTCTACGAGGCCATCAGGAAC 3552
QY 472 uValArgCysIleArg---PheAspAsnLysArgIleValSerGlyAlaTyrAspGlyLy 491
DB 3553 AGTGAAGACTTTAGACTCTTGA AAAATTCAGAGACTGCTTTCTTGGTCATTTGATGGAAC 3612
QY 491 sIleLysValTrpAspLeuValAlaAlaLeuAspProArgAlaProAlaGlyThrLeuCy 511
DB 3613 AGTGAAGTATGGAATATTTACTGGAATAAGAA----- 3649
QY 511 sLeuArgThrLeuValGluHisSerGlyArgValPheArg-----LeuGlnPheAspG1 529
DB 3650 ---AAAGACTTTCTCTGTCTCACCAGGTGACAGTACTTTCTTGTGACATTTCTCACCAGTGC 3705
QY 529 uPheGlnIleValSerSerHisAspAspThrIleLeuIleTrpAspPhe 546
DB 3706 TACCAAGTTTCTACTCTCTCTGTCGACAGACTGCAAGATCTGGAGTTT 3757

RESULT 12
US-09-690-364-21
; Sequence 21, Application US/09690364
; Patent No. 6468795
; GENERAL INFORMATION:
; APPLICANT: Hong Zhang
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF APAF-1 EXPRESSION
; FILE REFERENCE: RTS-0190
; CURRENT APPLICATION NUMBER: US/09/690,364
; CURRENT FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 100
; SEQ ID NO 21
; LENGTH: 7042
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (578)...(4162)
US-09-690-364-21

Alignment Scores:
Pred. No.: 1,59e-23 Length: 7042
Score: 299.50 Matches: 154
Percent Similarity: 33.77% Conservative: 102
Best Local Similarity: 20.32% Mismatches: 195
Query Match: 9.87% Indels: 308
DB: 4 Gaps: 34
```

US-09-601-168B-2 (1-569) x US-09-690-364-21 (1-7042)

```
Qy 2 AspProAlaGluAlaValLeuGlnGluLysPheMetAsnSerSerGluArg 21
Db 1769 GAAGAAGTTGAAGACATACACTGCTGAGAG-----TTTGTAAATAAGTCTCTTTTA 1816
Qy 22 GluAspCysAsnAsnGlyGluProArg-----Lysile 33
Db 1817 TTCTGTGATCGAATGGAAGTCTGTTTCGTTATTATTATCATGATCTTCAAGTAGATTTT 1876
Qy 34 IleProGluLysAsnSerLeuArg----- 41
Db 1877 CTTACAGAGAAGAAATTCAGCCAGCTTCAGGATCTACATAAGAAGATAATCACTCAGTTT 1936
Qy 42 GlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGlu-ThrValCysLeuAlaase 61
Db 1937 CAGACATATCACCAGCCCATCTCTTCCACCAGATCAGGAAGACTGTATGTTGGTAC 1996
Qy 61 rThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSerSe 81
Db 1997 AACTTT-CTGCCCTATCAC----- 2014
Qy 81 rMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeuCy 101
Db 2015 -----ATGCCAGTCCCAAGATGCACAGGAAGCTTTG 2046
Qy 101 sValLysTyrPheGlu-----GlnTyrSerGluSerAspGlnValGluPheVal-----G1 118
Db 2047 TGCTTTAATGTTTCCCTCGGATTCGGATTAAAGCA---AAACAGAACTTGTAGGCCCTGC 2103
Qy 118 uHisIleLeuSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysPr 138
Db 2104 TCATCTGATTTCATGATTTGGTATACAGACAT-----AT 2139
Qy 138 oMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaG1 158
Db 2140 ACTGATGAAAAGGATTGTGCA-----GTCAGTGA 2169
Qy 158 uAsnIleLeuSerTyrLeuAspAlaLys----- 167
Db 2170 GAATTTTCAGGAGTTTATCTTTAAATGGACACTTCTTGGACGACGCCATTTCTCTAA 2229
Qy 168 -----SerLeuCysAlaAlaGluLeuValCysLysGluTyrTyr----- 180
Db 2230 TATTGTACAACACTGGGTCTCTGTGAGCCGGA-----ACTTCAGAAGTTTATCAGCAAGC 2283
Qy 181 -----ArgValThrSerAspGlyMetLeuTyr-----Ly 190
Db 2284 TAAGCTGCAGGCCCAAGCAGGAGTCGATAATGGAAATGCTTTACTGGAATGGATAAACAA 2343
Qy 190 sLysLeuIleGluArgMetValArg-----ThrAspSerLeuTyrPr 204
Db 2344 AAAAACAATCATCGAATCTTTCCCGCTTAGTTGTCCGCCCCACACAGATGCTGTT----- 2398
Qy 204 gGlyLeuAlaGluArgArgGlyTyrGlnTyrLeuPheLysAsnLysProAspG1 224
Db 2398 ----- 2398
Qy 224 yAsnAlaProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleG1 244
Db 2399 -----TACCATGCTGCTGTTTCT-----GAGCATGCTCA 2427
Qy 244 uThrIleGluSerAsnTyrPArgCysGlyArgHisSerLeuGluArgIleHisCysArgSe 264
Db 2428 GAGATAGCTTCT-----TGTTGGAGCTGATAAAACCTTACAGGTGTTTC---AARGC 2475
Qy 264 rGluThrSerLys-----GlyValTyrCysLeuG1 274
Db 2476 TGAACAGGAGAGAACTTCAGAAATCAAGGCTCATGAGATGAAGTGTGTTGTTGTCG 2535
Qy 274 nTyr-----AspAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTr 292
Db 2536 ATTCTCTACAGATGACAGATTATTAGCAACCTGCTCATGTGATAAAAAAAGATGAAGATTG 2595
```



```

QY 511 sleuArgThrIleuValGluHisSerHisAspThrIleLeuIleTyrAspPhe 546
DB 3650 ----AAAGACTTCTCTCCAGGGTACAGTACTTTCTTGACATTTCTCAGCATGC 3705
QY 529 uPheGlnIleValSerSerHisAspThrIleLeuIleTyrAspPhe 546
DB 3706 TACCAAGTTTCACTCTCTGTCGACAAAGCTGCAAGACTCTGGAGTTT 3757

RESULT 13
US-09-092-508-15
; Sequence 15, Application US/09092508
; Patent No. 6291643
; GENERAL INFORMATION:
; APPLICANT: Henzel, William J.
; TITLE OF INVENTION: APAF-1, AN ACTIVATOR OF C ASPASE-3
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; STREET: 3100 No. 6291643west Center, 90 South Seventh St
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/092.508
; FILING DATE: 05-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/048,807
; FILING DATE: 05-JUN-1997
; APPLICATION NUMBER: 60/055,258.
; FILING DATE: 07-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Ph.D., Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 11669.6USU1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7075 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 578...4192
; OTHER INFORMATION:
US-09-092-508-15

Alignment Scores:
Pred. No.: 1.6e-23 Length: 7075
Score: 299.50 Matches: 154
Percent Similarity: 33.77% Conservative: 102
Best Local Similarity: 20.32% Mismatches: 195
Query Match: 9.87% Indels: 308
DB: 3 Gaps: 34

US-09-601-168B-2 (1-569) x US-09-092-508-15 (1-7075)

QY 2 AspProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGluArg 21
DB 1802 GAAGAAGTTGAAGACATCTCAGGAG-----TTGTAAATAAGTCTCTTTTA 1849
QY 22 GluAspCysAsnGlyGluProArg-----Lysile 33

```

```

DB 1850 TTCTGTGATCGGAATGGAAGTCTGTTTATTATTATTACATGATCTTCAAGTAGATTTT 1909
QY 34 IleProGluLysAsnSerLeuArg----- 41
DB 1910 CTTACAGAGAAGAAATTCAGCCAGCTTCAGGATCTACATAAGAGATAATCACTCAGTTT 1969
QY 42 GlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGlu-ThrValCysLeuAlaSe 61
DB 1970 CAGAGATATCCACCGCCGATCTTTTCCACCATCAGGAAGACTGTATGTATGTATC 2029
QY 61 rThrAlaMetLysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSerSe 81
DB 2030 AACTTT-CTGGCCTATCAC----- 2047
QY 81 rMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysGluLeuCy 101
DB 2048 -----ATGCCAGTGCCCAAGATGCCAAGAACTTTC 2079
QY 101 sValLysTyrPheGlu---GlnTyrSerGluSerAspGlnValGluPheVal-----G1 118
DB 2080 TGCCTTAATGTTTCCCTGGGATTGGATTAAAGCA---AAACAGAACTTTAGGCCCTGC 2136
QY 118 uHisLeuIleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysPr 138
DB 2137 TCATCTGATTCAATGAATTTGTGAATACAGACAT-----AT 2172
QY 138 oMetLeuGlnArgAspPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaG1 158
DB 2173 ACTAGTAAAGGATTGTGCA-----GTCAGTGA 2202
QY 158 uAsnIleLeuSerTyrLeuAspAlaLys----- 167
DB 2203 GAATTTTCAGGAGTTTATCTTTAAATGGACACCTTCTTGGACGACAGCCATTTCCTAA 2262
QY 168 -----SerLeuCysAlaAlaGluLeuValCysLysGluIleTyr----- 180
DB 2263 TATTGTACAACCTGGGTCTCTGAGCCGAA-----ACTCAGAACTTTATCAGCAAGC 2316
QY 181 -----ArgValThrSerAspGlyMetLeuTyr-----Ly 190
DB 2317 TAAGCTGCAGGCCAAGCAGAGGTGCGATATGGAATGCTTTACCTGGAATGATGAACAA 2376
QY 190 sLysLeuIleGluArgMetValArg-----ThrAspSerLeuTyrPr 204
DB 2377 AAAAAACATCAGAACTCTTCCCGTTAGTTGTCCGCCCCACACAGATGCTGT----- 2431
QY 204 gGlyLeuAlaGluArgArgGlyTyrGlyGlnTyrLeuPheLysAsnLysProAspGl 224
DB 2431 ----- 2431
QY 224 yAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleGl 244
DB 2432 -----TACCATGCTGCTTTCT-----GAGGATGTGCA 2460
QY 244 uThrIleGluSerAsnTyrArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSe 264
DB 2461 GAGAAATAGCTTCT-----TGTGGAGCTGATAAACCTTACAGGTGTTC---AAAGC 2508
QY 264 rGluThrSerLys-----GlyValTyrCysLeuGl 274
DB 2509 TGAACACGAGAGAAACTTCTAGAAATCAAGCTCATGAGGATGAAGTGTCTTTGTGTGC 2568
QY 274 nTyr-----AspAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTyr 292
DB 2569 ATTCTCTACAGATGACAGATTTATAGCAACCTGCTCAGTGTGATAAAAAGTGAAGATTG 2628
QY 292 pAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCy 312
DB 2629 GAATTTCTATGCTGGGAACCTAGTACACACCTATGATGAGCAGCTCAGAGCAAGTCAATTG 2688
QY 312 sLeuGlnTyrAspGluArg-----ValIleIleThrGlySerSerAspSerTh 328

```

```

Db      2689 CTGCCATTTTCAACCAAGTAGTCACTCTCTCTTAGCCACTGGGTCAAGTGACTGCTT 2748
Qy      328 rValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisCysG1 348
Db      2749 CCTCAAACTTTGGGATTTGAATCAAAAAGAAATGTCGAAATACCATGTTTGGTGCATACAAA 2808
Qy      348 uAlaValLeuHisLeuArgPhe-----AenAsnGlyMetMetValThrCysSerLysAs 366
Db      2809 TTCAGTCAATCACTGCAGATTTTACCAGATGATAGCTTTTGGCTAGTGTGTTCAGCTGA 2868
Qy      366 pArgSerIleAlaValTrpAspMetAlaSerProThrAsp-----IleThrLeuArg 383
Db      2869 TGGAACTTAAAGCTTTGGGATCGCATCAGCAAAATGAGAGAAAGCAATTAATGTGAA 2928
Qy      383 g----- 383
Db      2929 ACAGTTCTTCTTAATTTGAGGACCTCAGAGGATATGGAAGTGATAGTGAAGTGTG 2988
Qy      384 -----ArgValLeuVal----- 387
Db      2989 TTCGTGCTCTGCTGATGGTCAAGGATAATGGTGGCAGCAAAAATAAAATCTTTTGTG 3048
Qy      388 -----GlyHisArgAlaAlaValAsnVa 395
Db      3049 GAATACAGATCAGGTTCAAGGTGGCTGNTTCAGAGGACATTTAAGTGGTTCATGG 3108
Qy      395 lValAspPhe-----AspAspLysTyrIleValSerAlaSerGlyAspArgThrIleLy 413
Db      3109 TGTGATGTTTCTCTGATGATCATCAATTTTGTGACATCTTGTGATGACCAACAATCAG 3168
Qy      413 sValTrpAsnThr----- 417
Db      3169 GCTCTGGAGACAAAGAAAGTATGTAAAGACTCTGCTGTAAATGTTAAAGCAAGAAAGTAGA 3228
Qy      417 ----- 417
Db      3229 TGTGTGTTTCAAGAAATGAAGTGATGGTCTTGCAGTTGACCATATAGAAGCTGCA 3288
Qy      418 -----SerThrCys 420
Db      3289 ACTCATTAAATGGAAGAACAGGTGAGATTTATCTGACTGAAGTCAAGTTAGTGTG 3348
Qy      420 s-----GluPhe 422
Db      3349 TTGCTTAAGTCCACATCTTCAGTACATTTGATTTGGAGATGAATAATGGAGCCATGAGAT 3408
Qy      422 eValArgThrLeuAsnGly-----HisLysArgGlyIleAl 434
Db      3409 TTTAGAACTTGTAAACAAATAGAATCTTCAGTCCAGTTTCCAGCACAGAAACATGTATG 3468
Qy      434 aCysLeuGlnTyrArg-----AspArgLeuValValSerGlySerSerAspAsnThrIl 452
Db      3469 GCACATCCAGTTCACAGCCGATGAGAAGACTCTTATTCAAGTTCTGATGATGCTGAAAT 3528
Qy      452 eArgLeuTrpAspIleGluCysGlyAlaCysLeuArgValLeuGluGlyHisGluLe 472
Db      3529 TCAGGTATGAATTTGGCAATTTGCACAAATGTATC---TTTCTACGAGGGCCATCAGCAAC 3585
Qy      472 uValArgCysIleArg---PheAspAsnLysArgIleValSerGlyAlaValAspGlyLy 491
Db      3586 AGTGAAGACTTTAGACTCTTGAAATTTCAAGACTGCTTTCTTGTCTATTGATGGAAC 3645
Qy      491 eLysValTrpAspLeuValAlaLeuLeuAspProArgAlaProAlaGlyThrLeuCy 511
Db      3646 AGTGAAGTATGGAATATTTACTGGAATAAAGAA----- 3682
Qy      511 sLeuArgThrLeuValGluHisSerGlyArgValPheArg-----LeuGluPheAspGl 529
Db      3693 -----AAAGACTTTCTGTCACCGGGTACAGTCTTTCTTGTGACATTTCTCAGATGC 3738
Qy      529 uPheGlnIleValSerSerHisAspAspThrIleLeuIleTrpAspPhe 546
Db      3739 TACCAAGTTTTCATCTACTCTCTGTCAGAACAGTGCACAAAGATCTGGAGTTT 3790

```

```

RESULT 14
US-09-435-115-15
; Sequence 15, Application US/09435115
; Patent No. 6346607
; GENERAL INFORMATION:
; APPLICANT: Henzel, William J.
; TITLE OF INVENTION: APAF-1, AN ACTIVATOR OF C ASPASE-3
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Bdeil, Welter & Schmidt
; STREET: 3100 No. 6346607 West Center, 90 South Seventh St
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/435,115
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/092,508
; FILING DATE:
; APPLICATION NUMBER: 60/055,258
; FILING DATE: 07-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Ph.D., Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 11669, 60SU1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7075 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 578...4192
; OTHER INFORMATION:
; US-09-435-115-15
Alignment Scores:
Pred. No.: 1,6e-23 Length: 7075
Score: 299.50 Matches: 154
Percent Similarity: 33.77% Conservativeness: 102
Best Local Similarity: 20.32% Mismatches: 195
Query Match: 9.87% Indels: 308
DB: 4 Gaps: 34
US-09-601-168b-2 (1-569) x US-09-435-115-15 (1-7075)
Qy      2 AppProAlaGluAlaValLeuGlnGluLysAlaLeuLysPheMetAsnSerSerGluArg 21
Db      1802 GAAGAAGTTGAAGACATCTGAGGAG-----TTGTAAATAAGTCTCTTTTA 1849
Qy      22 GluAspCysAsnAsnGlyGluProArg-----Lysile 33
Db      1850 TTCTGTGATCGGAATGGAAGTCGTTTCGTTATTATTACATGATCTTCAAGTAGATTTT 1909
Qy      34 lleProGluLysAsnSerLeuArg----- 41
Db      1910 CTTACAGAGAGAAATTGCGAGCGCAGCTTCAGGATCTACATAGAAGATAATCACTCAGTTT 1969

```

QY 42 GlnThrTyrAsnSerCysAlaArgLeuCysLeuAsnGlnGlu-ThrValCysLeuAlaSe 61
 DB 1970 CAGAGATATCAACAGCCGACATACCTTTCCACAGATCAGGAAGACTGTATGATTCGTAC 2029
 QY 61 rThrAlaMetIysThrGluAsnCysValAlaLysThrLysLeuAlaAsnGlyThrSerSe 81
 DB 2030 AACITTT-CTGCCCTATCAC----- 2047
 QY 81 rMetIleValProLysGlnArgLysLeuSerAlaSerTyrGluLysGluLysLeuCy 101
 DB 2048 -----ATGCCAGTGCACCAAGATGCACCAAGCACTTTG 2079
 QY 101 svallystyrPheGlu--GlnTrpSerGluSerAspGlnValGluPheVal-----G1 118
 DB 2080 TGCTTTAATGTTTCCCTGGATTGGATTAAAGCA---AAACAGAACTGTAGGCCCTGC 2136
 QY 118 uHisLeuIleSerGlnMetCysHisTyrGlnHisGlyHisIleAsnSerTyrLeuLysPr 138
 DB 2137 TCATCTGATTCAATGAATTTGTGGAATACAGACAT-----AT 2172
 QY 138 oMetLeuGlnArgPheIleThrAlaLeuProAlaArgGlyLeuAspHisIleAlaG1 158
 DB 2173 ACTAGATGAAAAGGATTGTGCA-----GTCAGTGA 2202
 QY 158 uAsnIleLeuSerTyrLeuAspAlaLys----- 167
 DB 2203 GAATTTTCAGGAGTTTATCTTTAAATGGACACTCTTGGACGACAGCCATTTCTTAA 2262
 QY 168 -----SerLeuCysAlaAlaGluLeuValCysLysGluTrpTyr----- 180
 DB 2263 TATTGTCAACTGGCTCTCTGTGACCGGAA-----ACTTCAGAGTTTATCAGCAAGC 2316
 QY 181 -----ArgValThrSerAspGlyMetLeuTrp-----Ly 190
 DB 2317 TAAGCTGCAGCCCAAGCAGAGGTCGATATGGAATGCTTTACCTGGAAATGGATAACAA 2376
 QY 190 slsyleuIleGluArgMetValArg-----ThrAspSerLeuTrpPar 204
 DB 2377 AAAAAACATCAGAACTCTTTCCCGCTTAGTTGTCGCCGCCACACAGATGCTGTT----- 2431
 QY 204 gGlyLeuAlaGluArgArgGlyTyrGlyGlnTyrLeuPheLysAsnLysProProAspG1 224
 DB 2431 ----- 2431
 QY 224 yAsnAlaProProAsnSerPheTyrArgAlaLeuTyrProLysIleIleGlnAspIleG1 244
 DB 2432 -----TACCATGCCTGCTTTTCT-----GAGGATGGTCA 2460
 QY 244 uThrIleGluSerAsnTrpArgCysGlyArgHisSerLeuGlnArgIleHisCysArgSe 264
 DB 2461 GAGAATAGCTTCT-----TGTTGAGCTGATAAACCCTTACAGGTGTTCT---AAAGC 2508
 QY 264 rGluThrSerLys-----GlyValTyrCysLeuG1 274
 DB 2509 TGAACACAGAGAGAAACCTTCTAGAAATCAAGGCTCATGAGGATGAAGTGTGTTGTTGTC 2568
 QY 274 nTyr-----AspAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTr 292
 DB 2569 ATTCCTTACAGATGACGATTTATAGCAACCTGCTCAGTGGATAAAAAGTGAAGATTG 2628
 QY 292 pAspLysAsnThrLeuGluCysLysArgIleLeuThrGlyHisThrGlySerValLeuCy 312
 DB 2629 GAATTCATGACTGGGAACCTAGTACACACCTATGATGAGCACTCAGACGAAGTCAATTG 2688
 QY 312 sLeuGlnTyrAspGluArg-----ValIleIleThrGlySerSerAspSerTh 328
 DB 2689 CTGCCATTTTCAACACAGTAGTCATCTTCTTGTGCCACTGGTCAAGTGAAGTCTGCTT 2748
 QY 328 rValArgValTrpAspValAsnThrGlyGluMetLeuAsnThrLeuIleHisCysG1 348
 DB 2749 CTCAAACTTTTGGATTGAAATCAAAAAGATGTGAAAATACCATGTGTTGTCATACAAA 2808
 QY 348 uAlaValLeuHisLeuArgPhe-----AsnAsnGlyMetMetValThrCysSerLysAs 366

DB 2809 TTCTGTCATCACTGCGATTTTCCACAGATGAATAGCTTTTGGCTAGTGTGTTCACTGA 2868
 QY 366 pArgSerIleAlaValTrpAspMetAlaSerProThrAsp-----IleThrLeuArg 383
 DB 2869 TGGAACTTTAAAGCTTTGGGATCGGACATCAGCAAAATGAGAGGAAAAAGCATTAATGTGAA 2928
 QY 383 g----- 383
 DB 2929 ACAGTTCTTCTAAATTTGGAGACCCCTCAAGAGATATGGAAGTGATAGTGAAGTTTG 2988
 QY 384 -----ArgValLeuVal----- 387
 DB 2989 TTCTGCTGCTGCTGATGGTCAAGGATAATGTGGCAGCAAAAAATAAATCTTTTGTG 3048
 QY 388 -----GlyHisArgAlaAlaValAsnVa 395
 DB 3049 GAATACAGACTCAGTTCAAAGGTGGCTGATTCAGAGGACATTTAAGTTGGTTTCATCG 3108
 QY 395 lValAspPhe-----AspAspLysTyrIleValSerAlaSerGlyAspArgThrIleLy 413
 DB 3109 TGTGATGTTTCTCTGATGATGATCACTTTTGTGACATCTTCTGATGACGACACATCAG 3168
 QY 413 sValTrpAsnThr----- 417
 DB 3169 GCTCTGGGAGACAAAGAAAGATATGTAAGAACTCTGCTGTAATGTTAAAGCAAGATAGA 3228
 QY 417 ----- 417
 DB 3229 TGTGTTGTTCAAGAAAAATGAAGTATGTCCTTCGAGTTGACCATATAAGACGTCTGCA 3288
 QY 418 -----SerThrCy 420
 DB 3289 ACTCATTAATGAAGAAGACAGTCAGATGATGATATCTGAAGCTCAAGTTAGCTGCTG 3348
 QY 420 s-----GluPh 422
 DB 3349 TTGCTTAAGTCCACATCTTCAGTACATTCATGATTTGGAGATGAAATGGAGCCATTGAGAT 3408
 QY 422 eValArgThrLeuAsnGly-----HisLysArgGlyIleAl 434
 DB 3409 TTTAGAACTTTGTAACAATAAGAACTTCCAGTCCAGGTTTCAGCACAAAGAAACTGTATG 3468
 QY 434 aCysLeuGlnTyrArg-----AspArgLeuValValSerGlySerSerAspAsnThrI1 452
 DB 3469 GCATCCAGTTCACACCGATCAGAGACTCTTATTTCAAGTTCTGATGATGCTGAAAT 3528
 QY 452 eArgLeuTrpAspIleGluCysGlyAlaCysLeuArgValLeuGluGlyHisGluGluLe 472
 DB 3529 TCAGGTATGGAATGGCAATGGCAAAATCTATC---TTTCTACGAGGCCATCAGGAAC 3585
 QY 472 uValArgCysIleArg---PheAspAsnLysArgIleValSerGlyAlaTyrAspGlyLy 491
 DB 3586 AGTGAAGACTTTAGACTCTTGAAAAATTCAGACTGCTTTCTTGTGCTATTTGATGGAAC 3645
 QY 491 sIleLysValTrpAspLeuValAlaAlaLeuAspProArgAlaProAlaGlyThrLeuCy 511
 DB 3646 AGTGAAGTATGGAATATTTACTGGAATAAGAA----- 3682
 QY 511 sLeuArgThrLeuValGluHisSerGlyArgValPheArg-----LeuGlnPheAspG1 529
 DB 3683 ---AAAGACTTTCTGTCACCAAGGTACAGTACTTTCTTGTGACATTTCTCAGCATGC 3738
 QY 529 uPheGlnIleValSerSerHisAspAspThrIleLeuIleThrAspPhe 546
 DB 3739 TACCAAGTTTTCATCTACCTCTGCTGACAGACTGCAAGATCTGGAGTTTT 3790

RESULT 15

US-08-188-582-17
 ; Sequence 17, Application US/08188582
 ; Patent No. 5534410
 ; GENERAL INFORMATION:
 ; APPLICANT: Tjian, Robert

APPLICANT: Comai, Lucio
 APPLICANT: Dynlact, Brian D.
 APPLICANT: Roey, Timothy
 APPLICANT: Ruppert, Siegfried
 APPLICANT: Tanese, Naoko
 APPLICANT: Wang, Edith
 APPLICANT: Weinzierl, Robert O.J.
 TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
 TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
 NUMBER OF SEQUENCES: 36
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FLEHR, ROHBACH, TEST, ALBRITTON & HERBERT
 STREET: 4 Embarcadero Center, Suite 3400
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94111-4187
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/188,582
 FILING DATE: 28-JAN-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Osman, Richard A.
 REGISTRATION NUMBER: 36,627
 REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 781-1989
 TELEFAX: (415) 398-3249
 TELEX: 910 27299

INFORMATION FOR SEQ ID NO: 17:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2152 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:

NAME/KEY: CDS

LOCATION: 1..2112

US-08-188-582-17

Alignment Scores:

Seq. No.: 2,98e-24 Length: 2152
 Core: 298.50 Matches: 100
 Percent Similarity: 38.6% Conservative: 54
 Percent Local Similarity: 25.13% Mismatches: 143
 Query Match: 9.84% Indels: 101
 Gaps: 15

US-09-601-168b-2 (1-569) x US-08-188-582-17 (1-2152)

y 218 LysAsnLysProProAspGlyAsnAlaProProAsnSerPheTyrArgAlaLeuTyrPro 237
 |||||
 b 949 AAAAGCAAAAACAGATCCCAATGCTCCACCTCAGAAC-----AGAATCCCTCTTCT 1002

y 238 LysIleIleGlnAspIleGluThrIleGluSerAsnTrpArgCysGlyArgHisSerLeu 257
 |||||
 b 1003 GAG---TTGAAGATTTCAGAT-----AAGTTG 1026

y 258 GlnArgIleHisCysArgSerGluThrSerLys-----GlyValTyrCysLeu 273
 |||||
 b 1027 GATAAGATAATGAATATGAAAGAAACCAACCAACGAGTACGCTTGGCGGACTGCTTA 1086

y 274 GlnTyr----- 275

b 1087 CCCTCCATTCTTCTATACATTCTCAATGCTTACAGGGTCTCACTGAGTGGATGTC 1146

y 276 -----AspAspGlnLysIleValSerGlyLeuArgAspAsnThrIleLysIleTrp--- 292

Db 1147 ACTGATGTTCTAGTCTGATTGCTGGAGTTTTCAGATTCACTGTCAGAGTGTGTCG 1206
 Qy |||||
 Qy 292 |||||
 Db 1207 GTAACACCAAAAAGCTTCGTAGTGTCAACAACAGCATCAGATCTTAGTCTTATAGACAAA 1266
 Qy |||||
 Qy 293 |||||
 Db 1267 GAATCAGATGATCTTAGAAGAATCATGATCAGAAACAGCAAGTGAAGTCAAGATT 1326
 Qy |||||
 Qy 303 LeuThrGlyHisThrGlySerValLeuCysLeuGlnTyr-----AspGluArgValle 320
 |||||
 Db 1327 TTGTATGTCACAGTGGGCTGCTACGGAGCCAGCTTCAGTCGGATAGGAATCTCTG 1386
 Qy |||||
 Qy 321 IleThrGlySerSerAspSerThrValArgValTrpAspValAsnThrGlyGluMetLeu 340
 |||||
 Db 1387 CTTTCTCTTCAGAGGAGGAACTGTTAGATTGGAGCCTTCAAACATTTACTTGTGTTG 1446
 Qy |||||
 Qy 341 AsnThrLeuIleHisHisCysGluAlaValLeuHisLeuArgPheAsn-----AsnGly 358
 |||||
 Db 1447 GTGGGATATAAAGGACACAACTATCCAGTATGGGACACACAAATTTCTCCATATGGATAT 1506
 Qy |||||
 Qy 359 MetMetValThrCysSerLysAspArgSerIleAlaValTrp-----AspMetAlaser 376
 |||||
 Db 1507 TATTTGTGTCAGGGGGCCATGACCGAGTACGCTCGGCTCGGCTACAGACCCTATCAG 1566
 Qy |||||
 Qy 377 ProThrAspIleThrLeuArgValLeuValGlyHisArgAlaAlaValAsnValVal 396
 |||||
 Db 1567 CCTTTA-----AGAATATTTCCGGCCATCTTGTGTGATGTAATTTGATC 1611
 Qy |||||
 Qy 397 AspPhe-----AspAspLysTyrIleValSerAlaSerGlyAspArgThrIleLysVal 414
 |||||
 Db 1612 AGATTCATCCAAATTCATATGTTCTCTACGGCTCTCGACGACAACTGTCGGCTC 1671
 Qy |||||
 Qy 415 TrpAsnThrSerThrCysGluPheValArgThrLeuAsnGlyHisLysArgGlyIleAla 434
 |||||
 Db 1672 TGGGACGCTCCTGAATGTAACGTGTGAAGGATCTTCACTGGACACAAAGGGACCAATTCAT 1731
 Qy |||||
 Qy 435 CysLeuGlnTyr-----ArgAspArgLeuValValSerGlySerSerAspAsnThrIle 452
 |||||
 Db 1732 TCCTTGACATTTTCTCCCAATGGGAGATTCCTGCTACAGGACCAACAGATGGCAGAGTG 1791
 Qy |||||
 Qy 453 ArgLeuTrpAspIleGluCysGlyAlaCysLeuArgValLeuGluGlyHisGluGluLeu 472
 |||||
 Db 1792 CTTCTTTGGGATATTGGACATGGTTGATGTTGGAGAAATTTGGCATCAGGTTCATGGATAAT 1851
 Qy |||||
 Qy 473 ValArgCysIleArgPhe-----AspAsnLysArgIleValSerGlyAlaTyrAspGly 490
 |||||
 Db 1852 GTCTGTTCACTTAGGTTTAGTAGAGATGGTGAATTTTGGCATCAGGTTCATGGATAAT 1911
 Qy |||||
 Qy 491 LysIleLysValTrpAspLeuValAlaAlaLeuAspPro----- 503
 |||||
 Db 1912 ACAGTTCCGATTATGGGATGCTATCAAGCCCTTTGAAGATTTTAGAGCCGATGACTTTACT 1971
 Qy |||||
 Qy 504 -----ArgAlaProAlaGlyThr-----LeuCysLeuArgThrLeu 515
 |||||
 Db 1972 ACAGCCACTGGGCATATAAATTTACCTCAGAAATTCACAGGAGTTATTGTTGGGAACATAT 2031
 Qy |||||
 Qy 516 ValGluHisSerGlyArgValPheArgLeuGlnPheAspGluPheGlnIleVal 533
 |||||
 Db 2032 ATGACCAATATCAACACCGAGTTGTACACCTTCATTTTACTCGAAGAAACCTGGTT 2085

Search completed: October 22, 2003, 17:13:31

Job time : 149 secs